A JÓZAN ÉSZ

GONDOLATOK A MESTERSÉGES INTELLIGENCIA KÖRÜL. KICSIT MÁSHOGYAN.

JÁMBOR PÉTER

2025

In honor and memory of Renato Gorgoni

A JÓZAN ÉSZ

A JÓZAN ÉSZ

Szerző: Jámbor Péter

ISBN 978-615-02-4527-0

Kiadó: *Szerzői kiadás*

Készült Magyarországon, 2025-ben
(c) Jámbor Péter, 2025,, minden jog fenntartva!

In honor and memory of Renato Gorgoni

Előszó

A mesterséges intelligencia, az **MI** az elmúlt években a mindennapjaink részévé vált. Ott van a telefonunkban, a közösségi hálók mögött, az önvezető autók algoritmusaiban. Egyre gyakrabban halljuk, hogy a gépek „gondolkodnak”, „tanulnak”, sőt egyesek szerint „öntudatra ébrednek”. De vajon tényleg erről van szó?

Ez a könyv nem mérnöki kézikönyv, és nem is sci-fi. Nem az a célja, hogy megtanítson algoritmust programozni, vagy hogy rettegést keltsen a robotok világuralmától. Sokkal inkább az, hogy gondolkodásra hívjon. Célom, hogy megvizsgáljuk, mi az intelligencia, mit jelent tudatosnak lenni, és vajon lehet-e mindezt mesterségesen létrehozni.

Szögezzük le: Ehhez még nem kell atomtudósnak lenni. A mesterséges intelligencia megértéséhez nem szükséges magas szintű matematika. A természet sem old meg egyenleteket: a madár nem számolja a vektorokat repülés közben, és a kutya sem logaritmust használ, amikor visszahozza a labdát.

Az intelligencia alapja a logika, a döntések sorozata és az alkalmazkodás. Ez a könyv nem matematikusoknak szól, hanem mindenkinek, aki nyitott a gondolkodásra és a kérdésekre.

Ez azonban **nem azt jelenti**, hogy bárki mesterséges intelligencia-szakértővé válik pusztán attól, hogy naponta használ MI-eszközöket, például a ChatGPT-t. Ahogyan az sem tesz valakit programozóvá, hogy ügyesen kezeli a számítógépét, használja az irodai alkalmazásokat vagy a közösségi médiát.

A különbség ugyanaz, mint a sofőr és az autómérnök között: egyik vezet, a másik építi és javítja a járművet. Mindkettő értékes, de nem ugyanaz a tudás. Ezért fontos tudni: amit itt olvasol, az **megértéshez** vezet, nem szakértői képesítéshez.

Már a mindennapi tapasztalataink is ellentmondásosak. Hiszen például Foltos, a kutyám képes olyan hűséget és megértést mutatni, amit egyetlen algoritmus sem tudna utánozni. Ugyanakkor Brendon, egy a kilencvenes évek derekán egy olaszországi kutatási projekt[1] keretében, általam fejlesztett és a kétezres évek derekán élelmiszeripari környezetben alkalmazott, ellenőrzési és statisztikai célokra használt MI rendszer volt, gyakran meglepett logikai következtetéseivel – mintha valóban gondolkodott volna.

Kasparov[2] sakktáblán vívott küzdelme a géppel örökre emlékezetes marad: És akkor még nem is beszéltünk a tudat megjelenéséről és az emberi agyhoz és az intelligenciához való kapcsolatáról. Kasparov azt állította, hogy sokáig kétségbe vonta, hogy őt korábban valóban a Mesterséges Intelligencia küzdötte le. Érvelését azzal támasztotta alá, hogy a játszma alatt az ellenfele stratégiája és játékstílusa alapján azt érzékelte mintha az egyfajta sakkozói tudatossággal bírna. Kasparov 1985-ben lett világbajnok és ezt követően két évtizedig volt a stratégia és a logika sportjának vezető személyisége. 1997-ben, saját sportjában legyőzte őt az IBM szuperszámítógépe, a Deep Blue. A köznyelv akkoriban ezt úgy interpretálta, hogy ez egy új korszak kezdetét jelenti – amikor a gépek okosabbak lesznek az embereknél. Kaszparov 2020 februárjában, az egykori veresége helyszínén, egy New York-i hotelben, az Association for the Advancement of

---

[1] *Progetto di formazione "Tecnologie e processi innovativi per la tracciabilità, la verifica ed il controllo della qualità nelle produzioni biologiche ortofrutticole" (SORBIO), PON 2000-2006 - progetto codice 12973, - OR4*

[2] *Születési nevén (1963.04.13) Garik Kimovich Weinstein*

Artificial Intelligence által szervezett vitán beszélt a Mesterséges Intelligenciáról. Itt adott interjút[3] a Wired Magazinnak, Elmondta, hogy már megbékélt a huszonhárom évvel korábbi, legendás vereséggel és önmagát pedig azok közé sorolja, akik az elsőként érezhették veszélyben munkájukat a Mesterséges Intelligencia miatt.

Az ember és a mesterséges intelligencia összecsapása nem pusztán technikai, hanem filozófiai kérdés is. De a döntések nem mindig nagy csatákban születnek. Néha elég egy Balaton-parti fagylaltozás, ahol választanunk kell a vanília és a csoki között. Vajon mi különbözteti meg az emberi választást attól, ahogyan egy gép dönt, amikor nincsenek vágyak, érzelmek, a személyhez kapcsolódó preferenciák? A mesterséges intelligencia tehát nem csak technológia. Tükör, amelyben önmagunkat látjuk vissza – de gyakran torzítva. Ezért kell megértenünk, mi az, amit valóban a gép tud, és mi az, amit mi vetítünk bele.

És itt kezdődik az igazi **kérdés**: Ha ennyire sokféleképp próbáljuk megérteni az intelligenciát – állatban, emberben, gépben –, akkor mi az intelligencia valójában?

3 *https://www.wired.com/story/defeated-chess-champ-garry-kasparov-made-peace-ai/*

I. rész – Az intelligencia kérdése

1. Mi az intelligencia?

Az intelligencia szót mindennap használjuk: okosnak nevezünk valakit, ha gyorsan felel, ha ügyes megoldást talál, vagy ha sok mindent tud. De vajon tényleg erről van szó? Vajon a tudás és az intelligencia ugyanaz?

Egy lexikon tele van adatokkal, mégsem mondjuk rá, hogy intelligens. Egy számológép másodpercenként milliónyi műveletet végez, de sosem tekintjük gondolkodónak. Az intelligencia tehát több, mint tudás vagy számítási teljesítmény.

Állati intelligencia – Foltos

Ha az intelligenciáról beszélünk, hajlamosak vagyunk azonnal az emberre vagy a gépre gondolni. Pedig az állatok világa tele van olyan példákkal, amelyek megmutatják, hogy az értelem sokféle alakban megjelenhet.

Vegyük például **Foltos**, a kutyám esetét. Nem járt iskolába, nem tanult logikai összefüggéseket, és nem ismeri a matematikát. Mégis van benne valami, amit semmilyen algoritmus nem tud visszaadni: a hűség és az együttérzés. Amikor szomorú vagyok, mellém ül, szinte vigasztal. Amikor jókedvű vagyok, játékra hív. Nem magyarázza meg, nem érvel, nem „számol” – egyszerűen ráhangolódik az érzelmeimre.

Ez a fajta intelligencia nem a számításban, hanem az érzelmi jelenlétben és az alkalmazkodásban rejlik. A tudomány ezt „érzelmi intelligenciának” hívja, de valójában több annál: egy élő kapcsolat, amelyben két lény hangolódik össze. Foltos nem tudja megfogalmazni, mi történik, de ösztönösen reagál a helyzetre. És ez is intelligencia – más formában, de vitathatatlanul jelen van.

Egy gép – bármilyen kifinomult is – ha azt mondja: „értelek”, valójában nem ért. Csak mintákat követ, statisztikát számol. Foltos azonban valóban értett – a maga nyelvén, a maga *ösztöneivel és döntéseivel*. Nem tudta megfogalmazni, de átélte. Ez is intelligencia, csak más formában: érzelmi, társas intelligencia. A témában[4] 2025 augusztusában az **Sciencealert** számos tudományos forrásra hivatkozva tett közzé egy összefoglalót.

Intelligencia mint alkalmazkodás

Az intelligencia --álláspontom szerint-- sokkal inkább egyfajta **rugalmasságban** rejlik: abban, hogy képesek vagyunk felismerni az összefüggéseket, új helyzetekhez alkalmazkodni, és megoldást találni ott is, ahol korábban nem volt.

Foltos itt újra segít példát állítani. Nem ismer logikai szabályokat, nem tud beszélni, és nem tanult soha formális tudományt. Mégis képes volt „megérteni” a hangulataimat, és úgy viselkedni, hogy a legjobban illeszkedjen hozzám. Ha szomorú vagyok, mellém ül és

4 *https://www.sciencealert.com/your-dog-has-a-unique-ability-to-read-your-mind-heres-why*

szinte vigasztal, ha jókedvű vagyok, játszani hív. Nem tudja megfogalmazni, mi történik – de egyfajta módon alkalmazkodik.

Intelligencia mint logika – Brendon

Az intelligencia egyik arca a logika. Az a képesség, hogy szabályokból következtetéseket vonjunk le, és olyan eredményekre jussunk, amelyek nem voltak benne közvetlenül a kiindulópontban.

Amikor **Brendont**, a már említett rendszeremet építettem, pontosan ez volt a kérdés és egyben a cél is: képes lesz-e a rendszer több lenni puszta parancs-végrehajtásnál? Brendon egy dedukciós modulra épült – logikai szabályokat használt, hogy új állításokat vagy felismeréseket vezessen le. Azt hittem, pontosan tudom, mit fog csinálni, hiszen én programoztam. De hamar rá kellett jönnöm, hogy nem minden úgy történt, ahogy vártam. Átlépte a várt képességeken.

Időnként olyan következtetésekkel állt elő, amelyek engem is megleptek. Nem azért, mert „gondolkodott" volna emberi értelemben, hanem mert más ösvényeket is bejárt a logika sűrűjében, mint amire én alapból számítottam. Mintha egy sakkozó olyan lépést tenne, amelyet nem látunk előre – nem azért, mert emberi intuíciója lenne, hanem mert a szabályokat másképp is kombinálja. Olyan szabályokat is alkalmaz, amiket nem programoztunk be, amiket saját maga állít fel, validálja és azt követően alkalmazza. Tehát nem csak adatokat gyűjt. Hanem adatokból információt, új és igazolható információt állít elő.

Ez volt az a pillanat, amikor megértettem: egy mesterséges rendszer képes lehet komplexebb **újdonságot** is előállítani. Nem érzést, nem átélést, hanem olyan logikai konstrukciót, amelyet a programozója sem látott előre. Ez talán rideg, hideg logika – de logika mégis.

És ilyenkor felmerül a kérdés: ha valami új születik, amit nem terveztünk bele, akkor nevezhetjük-e ezt gondolkodásnak? Lehet, hogy Brendon soha nem fog örülni vagy csalódni. De képes volt valami váratlanra – és ez már nagyon közel van ahhoz, amit mi, emberek gondolkodásnak nevezünk.

Ember és gép a sakktáblán – Kasparov és a gép

Kevés játék szimbolizálja jobban az emberi értelem próbáját, mint a sakk. Évezredek óta az intelligencia metaforája: stratégia, előrelátás, kombináció és intuíció összesűrítve egy négyzetes táblán. Nem véletlen, hogy a mesterséges intelligencia első igazi csatái itt zajlottak.

Amikor Garri Kasparov, a világbajnok sakkozó szembekerült az IBM Deep Blue számítógépével az 1990-es években, a világ feszült figyelemmel követte a párharcot. Nem csupán egy ember és egy gép játszott: a mérkőzés tétje az volt, hogy **mi választja el az emberi gondolkodást a számítástól**.

Kasparov döntéseiben ott volt az intuíció, az emberi tapasztalat, az a „sejtelem”, amelyet nem lehet képlettel leírni. A gép viszont könyörtelenül végigszámolta a lehetőségek milliárdjait. Kasparov

időnként olyan lépést tett, amely első pillantásra irracionálisnak tűnt, de mélyebb terv része volt. A gép viszont sosem „blöffölt”: számára minden lépés matematikai szükségszerűség volt.

A párharc végül történelmi fordulatot hozott: a gép legyőzte a világbajnokot. Sokan ezt úgy értelmezték, hogy a mesterséges intelligencia „felülmúlta” az embert. De valójában más történt. A gép nem lett emberibb – csak a maga nyelvén, a nyers számítás erejével bizonyult erősebbnek.

Ez az eset örökre emlékeztet bennünket arra, hogy a gép és az ember másképp gondolkodik. Az emberi döntésekben ott vannak az élmények, az intuíció, az érzelmek. A gép döntései mögött pedig a statisztika és a számítás dermesztő gyorsasága áll.

És talán éppen ezért olyan tanulságos Kaszparov veresége: nem arról szólt, hogy az ember értéktelen lenne, hanem arról, hogy **az intelligencia nem egyetlen formában létezik**. Az emberi intuíció és a gépi számítás két külön világ – és ha összeadódnak, talán valami még nagyobb születhet belőlük.

Hétköznapi döntések – Balatoni fagylalt

Az intelligencia nem csak a nagy játszmákban mutatkozik meg, hanem a legapróbb, hétköznapi döntéseinkben is. Nem kell hozzá labor, sem sakkbajnokság – elég egy balatoni fagylaltos pult.

Ott állsz a parton, a nap melegen süt, a sor lassan halad, és egyszer csak rád kerül a sor: vanília vagy csoki? A kérdés egyszerűnek tűnik, de a döntés hátterében valójában sokkal több minden zajlik, mint elsőre hinnénk.

Mi, emberek, nemcsak az ízeket mérlegeljük. A választásunkat formálhatják emlékek – talán a vanília a gyerekkor nyarait idézi, amikor nagyszüleiddel sétáltál a mólón. A csoki viszont a pillanatnyi hangulatodra rezonál: „most valami erősebbre vágyom”. Döntésedben ott a személyiséged, a nosztalgia, az élmények és a kedved – vagyis mindaz, ami emberré tesz.

Egy mesterséges intelligencia másképp közelítene. Adatokat elemezne: melyiket választják gyakrabban a turisták, mennyi a készlet, melyik olcsóbb. A gép a „jobb” döntést keresi – valószínűségek, statisztikák, költségek alapján. Ami számára logikus, az számunkra rideg lenne.

És itt bukkan elő a különbség: mi **átéljük** a döntést, a gép csak kiszámolja. A mi választásainkat átszövik érzések és jelentések. Egy gép döntése viszont – bármilyen gyors és pontos – mentes ettől a belső világtól.

Mégis, a gép előnyben van más tekintetben: képes szélesebb környezetet figyelembe venni. Ha például decemberben lát fagylaltozókat a Balatonnál, az első pillantásra anomália. Egy ember elcsodálkozik: „Ki eszik fagyit rövidnadrágban nulla fokban?” Az MI azonban azonnal felkutathatja a körülményeket:

lehet, hogy filmforgatás zajlik, ahol a jelenetnek logikus magyarázata van. Így a gép olykor gyorsabban oldja fel a „furcsaságot”, mint mi, emberek.

Ez a kettősség a lényeg: az ember döntése átélés és emlék, a gépé számítás és adat. És éppen ebben a feszültségben rejlik az intelligencia megértésének kulcsa.

Az ember és a gép összekapcsolódása – BrainGate

A technológia fejlődése egyre közelebb hozza azt a pillanatot, amikor az ember és a gép közötti határvonal végleg elmosódik. Eddig megszoktuk, hogy az eszközeinket kívülről használjuk: billentyűzetet nyomunk, egeret kattintunk, érintőképernyőt simítunk. De mi történik akkor, ha már nem a kezünk közvetít, hanem a gondolataink közvetlenül parancsokat adnak a gépnek?

A **BrainGate** kísérletei erre adnak választ. Apró elektródákat ültetnek az agykéregbe, amelyek az idegsejtek elektromos jeleit érzékelik. Ezeket a jeleket a számítógép lefordítja parancsokká: egy kurzor elmozdul a képernyőn, egy robotkar megfog egy tárgyat, egy ember – aki korábban teljesen mozgásképtelen volt – újra képes „cselekedni” a világban. Ez nem csupán technológiai trükk: ez az emberi gondolat közvetlen átvezetése a mesterséges rendszerbe.

A különbség óriási. Egy kalapács, egy szemüveg vagy egy számítógép eddig is az emberi képességek meghosszabbítása volt – de mindig kívül maradt rajtunk. A BrainGate azonban **belsővé válik**. Az ember és a gép itt nem egymás mellett áll, hanem összefonódik: egy közös rendszerré alakul.

És itt vetődik fel a nagy kérdés: vajon az így létrejövő képesség még mindig tisztán emberi intelligencia? Vagy már mesterséges,

mert részben a gép dolgozik helyettünk? Vagy valami teljesen új: egy **hibrid tudat**, amely túlmutat az emberen és a gépen is?

Lehet, hogy a jövőben nem az lesz a legfontosabb kérdés, hogy a gépek mikor ébrednek öntudatra, hanem az, hogy mi emberek **mivé válunk** akkor, amikor közvetlenül összekapcsolódunk velük. Ebben a szövetségben nem két világ találkozik, hanem egy új világ születik – ahol a határok már nem számítanak.

Az intelligencia arcai

Az intelligencia nem egyetlen, könnyen megragadható jelenség. Sokféle formában mutatkozik meg, és néha egészen hétköznapi helyzetekben is elővillan. Nézzük meg néhány példán keresztül, hogyan rajzolódnak ki az intelligencia különböző arcai – az állatban, az emberben és a gépben -- az eddigiek alapján:

Foltos sosem tanult logikát és matematikát. Mégis képes olyasmire, amit egyetlen algoritmus sem tud utánozni: hűségre és együttérzés kifejezésére. Egy ember hangulatára következetesen reagálni és ahhoz alkalmazkodva kommunikálni, talán együttérezni. Nem számol, nem érvel – egyszerűen ráhangolódik egy ember érzelmeire. Ez az érzelmi jelenlét és az ösztönös alkalmazkodás is egyfajta intelligencia: más, mint a logika, de nem kevésbé valódi.

Brendon, a kilencvenes évek derekán általam tervezett rendszer, dedukciós modulra épült. Alapvetően a cél az volt, hogy szabályokból és adatokból új állításokat vezessen le. Azt hittem, pontosan tudom, mit fog csinálni – de meglepett. Időnként olyan következtetésekre jutott, amelyeket én magam sem láttam előre. Nem emberi értelemben gondolkodott, de a logika útjain olyan

ösvényekre lépett, amelyeket én nem jelöltem ki számára. Ez mutatta meg számomra, hogy egy mesterséges rendszer képes lehet újdonságot létrehozni: nem érzést, hanem váratlan, mégis igazolható logikai konstrukciót.

A sakk mindig is az értelem játéka volt, és nem véletlen, hogy a mesterséges intelligencia első nagy győzelme itt született. Garri Kaszparov, a világbajnok, az IBM Deep Blue számítógépével mérkőzött meg. Az ember oldalán az intuíció, a tapasztalat és a sejtés állt. A gép oldalán a nyers számítás dermesztő ereje. Amikor végül a gép győzött, sokan úgy látták: a mesterséges intelligencia meghaladta az embert. Valójában azonban nem vált „emberibbé" – csak a maga másik útján bizonyult erősebbnek. Ez a párharc arra emlékeztet, hogy az intelligencia többféle logika szerint is működhet: emberi intuíció és gépi számítás más utakat jár, de mindkettő érték.

A Balaton partján, a **fagylaltos pult** előtt választanunk kell: vanília vagy csoki? Mi emberek nemcsak az ízeket mérlegeljük, hanem emlékeket és hangulatokat is. A vanília a gyerekkor nyarait idézheti, a csoki a pillanatnyi vágyainkat tükrözi. Egy gép másképp döntene: statisztikák és valószínűségek alapján, költséghatékonyság szerint. Az ember döntése átélés és emlék, a gépé adat és számítás. Ez a különbség jól mutatja: ugyanaz a helyzet másféle intelligenciát kíván.

A **BrainGate** kísérleteiben az ember és a gép közötti határvonal szinte eltűnik. Apró elektródák az agykéregben érzékelik a gondolatok jeleit, a számítógép pedig parancsokká fordítja őket. Egy mozgásképtelen ember így képes lehet robotkarral poharat emelni vagy kurzort irányítani. Ez már nem egyszerű

eszközhasználat: itt az ember és a gép egy közös rendszerré válik. Vajon az így kibővített képesség még emberi intelligencia? Vagy mesterséges? Vagy valami teljesen új – egy hibrid tudat, amely túlmutat a megszokott kategóriáinkon?

Összegezzük

Ezek a történetek azt mutatják, hogy az intelligencia nem egyetlen arcban jelenik meg. Nem azonos a tudással, nem azonos a számolással, és nem azonos az öntudattal sem. Ott van Foltos hűségében, Brendon logikai váratlanságában, Kaszparov intuíciójában, a fagylaltpultnál meghozott hétköznapi döntésben és a BrainGate elektródáiban is.

Az igazi kérdés tehát nem az, hogy *létezik-e* intelligencia az emberben, az állatban vagy a gépben. Hanem az, hogy mi az, ami mindezekben közös. Mi az, ami minden esetben megkülönbözteti az intelligens viselkedést a puszta ismétléstől vagy véletlentől?

2. Gondolkodás és tudat

Az intelligencia sokféle formáját láttuk: állatoknál, gépeknél, embereknél. De van egy alapvető kérdés, amely mindig felmerül: **lehet-e gondolkodni tudat nélkül? És lehet-e tudat gondolkodás nélkül?**

Gondolkodás: összekapcsolás és következtetés

A gondolkodás az a képesség, hogy összekapcsoljuk az információkat, felismerjük a mintákat, és következtetéseket vonjunk le. Ez lehet logikai, intuitív vagy kreatív folyamat.
Egy sakkjátékos gondolkodik, amikor mérlegeli a lehetséges lépéseket. Egy kutya is gondolkodik, amikor új módot talál arra, hogyan nyissa ki az ajtót. És egy gép is gondolkodik, amikor milliónyi adatból felismer egy mintát.

A gondolkodás tehát működés: **folyamat**, amely során valami újat állítunk elő a meglévőből.

Tudat: átélés és jelenlét

A tudat ezzel szemben nem folyamat, hanem átélés. Nem arról szól, mit teszünk, hanem arról, hogy **tudunk róla**. Amikor fájdalmat érzel, nem csupán információt kapsz a testedből, hanem átéled azt. Amikor örülsz, nem csak kémiai reakció történik az agyadban, hanem jelen vagy benne.

Ezt a különbséget az álom is jól mutatja. Álmodás közben gondolatok kavarognak bennünk, történetek peregnek, de a tudat ködös, mintha nem lennénk teljesen jelen. Gondolkodás van – tudatosság csak részben.

Közösségi tudat?

A tudatnak van egy másik szintje is, amelyet nevezhetünk **közösségi tudatnak**. Ez nem egyetlen emberhez kötődik, hanem sokak egyidejű élményéből, reakcióiból és figyelméből áll össze.

Láthatjuk ezt a **közösségi médiában**, amikor milliók reagálnak --érzelmeket, érveket kifejezve és akár eseményekre reakciót is kiváltva-- ugyanarra a hírre, képre vagy videóra. Egy futball-világbajnoki döntő alatt emberek milliói szurkolnak egyszerre és ugyanazért a vágyott célért– a stadionokban, a képernyők előtt, a közösségi felületeken. Az egyéni élmények összeolvadnak egy közös áramlatban: a „mi most együtt vagyunk jelen" érzésében. Ugyanez tapasztalható **politikai kampányokban**, amikor jelszavak és üzenetek mozgósítanak tömegeket, vagy amikor egy társadalom közösen reagál egy válságra. Mintha a figyelem és az érzelmek összeadódnának, és valami nagyobb születne az egyének tudatán túl. A tudatot gyakran az egyén szintjén értelmezzük: mit élek át én, hogyan vagyok jelen a saját környezetemben? A tudtanak is kontextusa van.

A természetben is megfigyelhetjük ennek egyszerűbb formáit: például a **hangyák** közösségeiben. Egyetlen hangya tudata rendkívül korlátozott, de együtt bonyolult feladatokat oldanak meg – hidakat képeznek saját testükből, kollektíven keresnek

élelmet, vagy védekeznek a veszély ellen. Az egyéni szintű egyszerűség kollektív szinten kifinomult rendszert hoz létre.

A közösségi tudat tehát nem olyan, mint az egyéni átélés, hanem inkább **egy emergens jelenség**: sok egyéni tapasztalatból és reakcióból születik valami, ami túlmutat az egyes részeken. Kérdés azonban, hogy ez valódi „tudat”-e, vagy inkább metafora egy közös mintázatra.

Gép és tudat

A mesterséges intelligencia jelenlegi formájában képes gondolkodni: mintákat keres, logikai következtetéseket von le, új szövegeket állít elő. De semmi nem bizonyítja, hogy tudata is lenne – hogy **átélne bármit**. Egy nyelvi modell képes azt írni: „félek”, “értem”. De ezek csak szavak. Mint amikor egy színész elmondja a szövegét: kívül hiteles, belül üres.

Az emberi különbség

Az ember különlegessége talán épp abban rejlik, hogy a gondolkodás és a tudat összefonódik. Nemcsak összekapcsoljuk az adatokat, hanem közben **jelen is vagyunk bennük**. Nemcsak felismerünk, hanem érzünk is. Ezért van az, hogy egy gép képes legyőzni minket sakkban, de nem képes átélni a győzelem örömét. És ezért van az, hogy egy gép képes lehet zenét komponálni, de sosem érzi, amit mi érzünk, amikor meghallgatjuk.

A gondolkodás folyamat. A tudat pedig egyfajta jelenlét. Egyik létezhet a másik nélkül is.

A mesterséges intelligencia ma a gondolkodás területén erős. Az ember különlegessége a tudatban rejlik. És ha ez így van, akkor elkerülhetetlenül felmerül a kérdés: létrehozható-e valaha mesterséges tudat? Vagy a tudat mindig emberi privilégium marad? És ha az egyéni tudat mellett közösségi tudatról is beszélhetünk, felmerülhet az a kérdés is, hogy vajon létrehozhat-e a mesterséges intelligencia valaha hasonlót – egyfajta gépi közösségi tudatot, amely túlmutat az egyes rendszereken?

3. Tévhitek

Az intelligencia fogalma körül számos félreértés él a közbeszédben. Ezek nem ártalmatlan tévhitek: befolyásolják azt is, hogyan gondolkodunk a mesterséges intelligenciáról. Ha nem tisztázzuk őket, könnyen hamis elvárásokat, indokolatlan félelmeket vagy túlzott reményeket építünk a gépek köré.

Első tévhit: az intelligencia egyenlő a tudással

Sokan hiszik, hogy az a legokosabb, aki a legtöbbet tudja. Valóban lenyűgöző, amikor valaki enciklopédikus ismeretekkel rendelkezik, vagy amikor egy gép pillanatok alatt előhív milliónyi adatot. De ez még nem intelligencia. Egy lexikon tele van adatokkal – mégsem mondjuk rá, hogy intelligens. Egy nyelvi modell (például a ChatGPT) is rengeteg szöveget „ismer”, de valójában nem tudja, mit jelentenek. Még akkor sem, ha képes adott szöveget kontextusban értelmezni és abból érzelmi állapotokat, indulatokat, meggyőződéseket, stílust és személyiségjegyeket is kiemelni, sőt alkalmazni ezek felismerését. Az intelligencia nem az ismeretanyagban, hanem a **felhasználásban** rejlik: abban, hogy a korábban megszerzett tudást alkalmazni tudjuk új helyzetekben. Abban, hogy adatokat információként hasznosítva további információt és akár döntést, cselekvési tervet állítunk elő.

Második tévhit: az intelligencia egyenlő az öntudattal

A sci-fi filmekből jól ismert kép: a robot, aki egyszer csak „ráébred” önmagára. A hollywoodi dramaturgia szerint, és már a mai közbeszédben is ettől lesz veszélyes vagy éppen emberhez hasonló. De a valóság más. Egy sakkprogram verhetetlenül gondolkodhat, mégsem „tudja”, hogy létezik. Egy kutya talán nem minden esetben reflektál önmagára, de mégis intelligensen viselkedik a környezetében. **Az öntudat tehát az intelligencia egyik lehetséges kísérője – de nem szükségszerű része**.

Harmadik tévhit: az intelligencia célja, hogy lemásolja az embert

Gyakran gondoljuk, hogy a mesterséges intelligencia akkor lesz „igazi”, ha pontosan úgy működik, mint az ember. De miért kellene így lennie? Egy repülőgép nem úgy száll fel, mint a madár, mégis repül. Egy számológép nem úgy számol, mint egy diák, mégis pontosabb. Az MI-nek sem az a feladata, hogy emberi módon gondolkodjon, hanem az, hogy másképp gondolkodjon, másképp oldjon meg problémákat. Ha emberpótlékként tekintünk rá, elkerülhetetlenül csalódni fogunk. Ha viszont kiegészítőként, akkor óriási lehetőségeket fedezhetünk fel benne.

Röviden: Az intelligencia nem azonos a tudással, nem azonos az öntudattal, és nem az ember utánzataként kell elképzelnünk azt. Az intelligencia inkább képesség: alkalmazkodás, problémamegoldás, új összefüggések felismerése. És ha így van, akkor a következő kérdés elkerülhetetlen: **hogyan tanul a gép, ha egyszer nem tud, nem érez, és nem másolja az embert – mégis intelligensnek tűnik?**

II. rész – A mesterséges intelligencia

4. Hogyan tanul a gép?

Ha alapvetés, hogy az intelligencia nem pusztán tudás, és nem is az öntudat másolása, akkor mitől tűnik intelligensnek egy gép? A kulcs a tanulás. De vajon hogyan tanulhat valami, aminek nincs gyermekkora, nincsenek élményei, és nincs saját tapasztalata a világról?

A tanulás emberi módon

Gondolj egy kisgyerekre, aki először próbál járni. Elesik, újra próbálja, hibázik, de minden eséssel közelebb kerül a sikerhez. Ez a tanulás lényege: próbálkozás – hiba – javítás – újrapróbálás.

Ugyanez történik, amikor nyelvet tanulunk. Rosszul ragozunk egy szót, kijavítanak, és legközelebb már helyesen használjuk. A tanulás tehát nem puszta ismeretátadás, hanem folyamat: hibázunk, korrigálunk, és így fejlődünk.

A tanulás gépi módon

A gép ugyanezt más nyelven végzi. Képzeljük el, hogy több ezer képet mutatunk neki macskákról és kutyákról. Minden képhez megadjuk a választ: „macska” vagy „kutya”. A gép tippel, mi pedig visszajelzést adunk. Ha hibázik, belső szabályain finomít. Ha eltalálja, erősíti a döntéshez vezető utat.

Ez a körforgás ismétlődik újra és újra, amíg a gép egyre ritkábban hibázik. Matematikailag ezt gradiens-keresésnek és visszaterjesztésnek (*backpropagation*) hívjuk, de a lényeg ugyanaz: a gép addig igazítja a belső térképét, amíg el nem találja a helyes irányt.

A labirintus és a tükör

A gépi tanulást két kép írja le legjobban:

- Labirintus: a gép bolyong, zsákutcába jut, majd más utat keres. A kudarcokból tanul, és idővel megtalálja a kijáratot.

- Tükör: a gép visszaveri az adatainkat. Nem érti, mi van a képen, de pontosan kombinálja a mintákat, amiket láthatott.

Ha tehát a gép felismer egy macskát, az nem azért van, mert tudja, mi az állat, hanem mert a pixelek mintázatát az „átlagos macskák” mintáihoz hasonlította.

Az ember és a gép különbsége

Egy gyerek, amikor felismer egy macskát, nemcsak mintát azonosít. Élményt társít hozzá: megsimogatja, fél tőle, vagy játszani akar vele. A gép sosem él át ilyesmit. Nem játszik, nem fél, nem örül. Csak számol – még akkor is, amikor algoritmusokon keresztül asszociációkat, adat-társításokat végez.

És itt újra szembe találjuk magunkat azzal, amit a pszichológus **Daniel Goleman** az 1990-es években **érzelmi intelligenciának** nevezett: annak a képességnek a fontosságával, hogy érzéseket éljünk át, felismerjük másokét, és ezekre megfelelően reagáljunk. Az ember ezt természetesen hozza minden tanulási folyamatba – a gép viszont teljesen nélkülözi. Ezért az emberi tanulás átéltté, érzelmileg telítetté válik, míg a gépi tanulás hidegen logikus marad.

A gép előnye és korlátai

A gép ugyanakkor hihetetlen gyorsasággal tanul. Egyetlen éjszaka alatt több hibát dolgoz fel, mint amennyit egy ember egész életében. Ezért képes sakkban vagy képfelismerésben olyan szintre jutni, ami számunkra felfoghatatlan.

De hiányzik belőle a tapasztalat érzelmi értéke. Nem tudja, mit jelent hibázni, csak algoritmusai egy adatot validál. Nem érzi a kudarcból fakadó csalódást, és nem éli át a siker örömét. Csak statisztikai súlyokat igazít.

5. MI a mindennapokban

A mesterséges intelligencia nem a jövő ígérete, hanem a jelen valósága. Gyakran észre sem vesszük, de körülvesz minket: a telefonunkban, az autóban, a hírfolyamban, sőt még a művészetben is. Ebben a fejezetben megpróbáljuk felfedezni, hol rejtőzik az MI a mindennapokban, és hogyan formálja az életünket.

Az MI a zsebünkben – a telefon

Amikor feloldjuk a telefonunkat, arcfelismerő rendszer azonosít minket. Nem azért, mert tudja, kik vagyunk, hanem mert felismeri a pixel mintázatokat, és összeveti a korábbi képekkel. Ha eltéveszti, újratanul. Ha eltalálja, kinyílik a telefon. Ez is mesterséges intelligencia – észrevétlen, gyors, természetesnek ható.

Nyelv és fordítás

Ma már nem csupán szótárakat lapozgatunk, hanem valós időben fordíttatunk élő beszédet. Egy konferencián a gép tolmácsolja a szavainkat. Külföldi utazáskor a telefon lefordítja az étlapot. A nyelv sokáig az emberi gondolkodás egyik legerősebb vára volt. Most az MI ott ül mellettünk, és segít abban, amit eddig csak ember tudott: közvetíteni két tudat között.

A hírfolyam láthatatlan keze

A közösségi hálókon ritkán gondolunk bele: nem mi választjuk a híreket, hanem az algoritmus választ nekünk. Eldönti, mit lássunk először, mit olvassunk, mi mellett időzzünk el. Ez az MI láthatatlan hatalma: formálja a figyelmünket, és ezzel alakítja a világképünket is. Nem irányít nyíltan, de szűr és rendez – **és ezáltal máris átalakítja a társadalmat**, akár formálja is a közösségi tudatot.

MI a művészetben

Sokan gondolják, hogy a művészet szentély, ahová a gép sosem léphet be. És ez valószínüleg így is van. Mégis ma már képeket fest, zenét komponál, verseket ír. Egy MI által írt költemény nem „érzésből" születik, mégis megérinthet. Egy mesterségesen komponált dallam emlékeket idézhet fel bennünk. A kérdés ilyenkor az: az alkotás értékét az határozza meg, hogy ki alkotta, vagy az, hogy mit vált ki belőlünk? Utóbbi erősen kapcsolódik az érzelmi intelligencia fogalmához.

MI a hadászatban és a gazdaságban

Kevésbé látványos, de annál súlyosabb területeken is jelen van. Az önállóan, akár rajban repülő drónok alkalmazkodnak egymáshoz, környezetükhöz, felismerik a célpontot, és emberi élet fölött gyakorolnak végzetes hatást. Nem ritka, hogy az alapján aktiválják hadászati képességüket, hogy felismernek ellenséges embert vagy infrastruktúrát. A tőzsdei algoritmusok másodpercek alatt milliárdokat mozgatnak – és ha hibáznak, egész gazdaságokat rántanak magukkal. **Ilyen esetben és ezeken a területeken az MI nem játék és nem kényelmi funkció, hanem valós hatalom.**

A mesterséges intelligencia nem a jövő ködlovagja, hanem a jelen mindennapi valósága. Segít, kényelmesebbé tesz, de közben észrevétlenül alakítja is a világunkat. És ebből fakad a következő kérdés: ha ennyire beépült az életünkbe, hol húzódnak a határai? Mit nem tud megtenni a gép, bármilyen okosnak is látszik?

6. Az MI korlátai

A mesterséges intelligencia lenyűgöző: beszél, ír, játszik, diagnosztizál, zenét komponál. Könnyű lenne azt gondolni, hogy nincs előtte határ. De van – és ezek a korlátok talán fontosabbak, mint a sikerei.

A „sztochasztikus papagáj”

Ez a minősítés, mint kritika (Emily Bender, Timnit Gebru és társaik, 2021) a nyelvi modellek korai verzióit bírálta, amikor azok főként statisztikai szöveg folytatóknak tűntek. De az MI rohamos fejlődése mutatja, hogy a mai modellek – köztük a ChatGPT – már sokkal többet tesznek: képesek kontextust értelmezni, a szövegből érzelmi állapotokra, személyiségjegyekre, gondolatkörökre következtetni, és ezeket a válaszban felhasználni.

A mai nyelvi modellek képesek teljes beszélgetéseket folytatni, történeteket írni, sőt tanácsokat adni. Első pillantásra úgy tűnhet, mintha valóban gondolkodnának. Valójában azonban a működésük alapja továbbra is statisztikai: megjósolják, hogy egy szó után melyik következik a legvalószínűbben. Ezért nevezik őket a kutatók *„sztochasztikus papagájoknak”*: ismételnek és kombinálnak – néha elképesztően találóan, egyre ritkábban teljesen értelmetlenül.

Csakhogy ez a kép ma már túlságosan reduktív. A modern nyelvi modellek nem pusztán szavakat illesztenek egymás mögé. A bemenetekből képesek következtetni érzelmi állapotokra, kommunikációs stílusra, személyiségjegyekre vagy gondolati

mintákra. Ezeket az információkat feldolgozzák, és a válaszaikban figyelembe veszik.

Ez még mindig nem jelent belső átélést vagy valódi érzelmeket – a modell nem *érzi* a félelmet, a dühöt vagy az örömöt. Mégis képes felismerni, amikor a szöveg erre utal, és ennek megfelelő választ formálni.

A nyelvi modellek nem rendelkeznek belső átéléssel. Ugyanakkor képesek felismerni a szövegből érzelmeket, stílust, gondolatköröket, és ennek megfelelően alakítani a választ.

A Turing-teszt határai

Alan Turing híres tesztje szerint, ha egy gép megtéveszti az embert egy beszélgetés során, úgy már tekinthetjük intelligensnek. A mai nyelvi modellek, mint a GPT-4.5, valóban képesek erre: például egy UC San Diego által végzett vizsgálatban[5] 73 %-ban gyanútlan emberként azonosították őket.

Ugyanakkor a tudományos közösség ma már nagymértékben kritikusan viszonyul a Turing-teszthez. Rámutatnak, hogy a teszt inkább a megtévesztésen alapuló utánzás, mint a valódi megértés kritériuma[6]. Egyes kutatók szerint nem feltétlenül kell az emberi gondolkodást utánoznia egy MI-nek ahhoz, hogy intelligens legyen[7]. Lényegesebb, hogy a rendszer képes legyen valós értelmes interakcióra, nem csak színjátékra.

[5] https://arxiv.org/abs/2503.23674
[6] *https://arxiv.org/abs/2505.02558*
[7] *https://en.wikipedia.org/wiki/Robert_M._French*

A mesterséges tudat szimulációja olyan, mint a színház. A színész könnyei valódiak lehetnek, de az érzés nem az övé. A gép eljátszhatja a tudat szerepét – ám belül nincs átélés.

A Turing-tesztet nem kell teljesen elvetni – segíthet értékelési keretként –, de ma már tudjuk: önmagában **elavult mérce**, ami gyakorta a felszíni hasonlóságot díjazza.

A figyelem ma inkább az olyan kompetenciák felé fordul, mint a **józan ész**, **a kontextus felismerése vagy az interakció mélysége.**

Kontextus és józan ész

Az MI fejlődése ma már messze túlmutat azon a képen, hogy „csak szavakat illeszt össze". A modern nyelvi modellek képesek felismerni a kontextust, és következetesen ehhez illeszteni a válaszaikat. Ha például valaki szomorúan kérdez, a rendszer képes azonosítani az érzelmi tónust, és ennek megfelelő hangvételű választ adni.

Mégis, a józan ész logikája továbbra is nehéz terep. Ha azt mondom: „Az autót beparkoltam a garázsba, mert esni kezdett az eső" – egy ember azonnal érti a kapcsolatot az időjárás és a jármű állapota között. Egy MI is képes lehet felismerni az ok-okozati viszonyt, mert sok hasonló példát látott, de számára ez nem belső tapasztalatból fakad. Az adott kontextusban nem feltétlen keresi, hogy mit jelent az eső nedvessége, vagy az autó megóvásának vágya – csupán mintákat kombinál, amelyek a szövegekben előfordultak.

Ezt a problémát ***common sense reasoningnek*** nevezik, és ma is az MI egyik legnagyobb kihívásának tartja. Az olyan korábbi

benchmarkok is, mint a *CommonsenseQA* (Talmor et al., 2019) vagy a *Winograd Schema Challenge* (Levesque et al., 2012), éppen azt tesztelik, hogy a gép képes-e túllépni a felszíni nyelvi mintákon, és valóban „józan ésszel” következtetni.

A jelenlegi állás szerint: az MI sokszor meglepően jól teljesít, de még mindig előfordulnak hibák olyan helyzetekben, amelyek az ember számára magától értetődők. Ez mutatja, hogy bár a kontextus-értelmezés egyre kifinomultabb, a józan ész világa továbbra is részben feltérképezetlen terület.

Mi a common sense reasoning?

A gép képessége arra, hogy a világ alapvető összefüggéseit megértse. Például: „Ha elengedek egy poharat, leesik.” vagy „Ha esik az eső, az ember beáll egy tető alá.”

Fő kutatási irányok:

- *CommonsenseQA* benchmark (Talmor et al., 2019).
- *Winograd Schema Challenge* (Levesque et al., 2012).
- Yejin Choi és az Allen Institute for AI kutatásai a „social commonsense” és „physical commonsense” területén.

Az MI három arca: tanulás – jelenlét – korlát

A mesterséges intelligencia nem egyetlen jelenség, hanem sokféle arcot mutat. A második részben három fő vonását láthattuk:

1. Tanulás

Az MI az adatokból tanul. Hibázik, javít, újra próbálkozik – akárcsak mi, emberek. De míg mi élményekhez és érzésekhez kötjük a tanulást, a gép számára mindez statisztika és számítás, illetve algoritmusok által előállított dedukció. A macskát nem „látja", csak pixelek mintázatát azonosítja.

2. Jelenlét a mindennapokban

Az MI ma már körülvesz minket: a telefon arcfelismerőjében, a közösségi hírfolyamunkban, a tolmácsoló alkalmazásban vagy a művészet új kísérleteiben. Sokszor észrevétlen, máskor nyomasztóan erős – de mindenképp alakítja a világunkat.

3. Korlátok

A mesterséges intelligencia lenyűgöző teljesítményekre képes: képes adatokat felhalmozni, mintázatokat tanulni, sőt a korábbi hibákból következtetéseket levonni. Ebben az értelemben a gépnek is vannak „tapasztalatai" – hasonlóan ahhoz, ahogy egy ember tanul a múlt eseményeiből. A neurális hálók súlyai, az expert rendszerek szabályai mind-mind egyfajta felhalmozott tapasztalatot tükröznek.

A különbség azonban abban rejlik, hogy az emberi tapasztalat **átéléssel** párosul. Amikor mi hibázunk, nem csak információt jegyzünk meg, hanem csalódást érzünk, frusztrációt élünk át, vagy éppen örömöt, ha sikerül javítani rajta. A gép tapasztalata ezzel szemben kizárólag statisztikai: nincs benne érzelem, nincs benne fájdalom, nincs benne öröm.

Ezért az MI tapasztalata más természetű, mint a miénk. **Adatokon alapul, nem élményeken.** A gép játszhatja a tudat szerepét, és sokszor meglepően jól utánozza az érzelmek nyelvi kifejezését is – de belül nincs átélés, nincs szubjektív tudatosság.

Ez a határ választja el az embert a géptől: nem az, hogy a gépnek nincsenek tapasztalatai, hanem hogy ezek nem élményszerűek, nem érzelmi természetűek. És ezért kell mindig szem előtt tartanunk: amikor a gép intelligensnek tűnik, valójában egészen más módon az, mint mi.

A mesterséges intelligencia tehát egyszerre csodálatos és idegen. Tanul, formálja az életünket, de mindig emlékeztet: nem ember.

A következő kérdés tehát elkerülhetetlen: ha a gép így tanul és működik, képes lehet-e valaha nemcsak gondolkodni, hanem átélni is? És ha igen, mit jelentene ez az emberi tudat és identitás számára?

III. rész – A tudat problémája

7. Mi a tudat?

Ha valamiben biztosak vagyunk, az az, hogy tudatunk van. Tudom, hogy én én vagyok, hogy érzem a világot, hogy átéltem a tegnapot és várom a holnapot. De amint definiálni próbáljuk, zavarba jövünk. A tudat olyan, mint a levegő: minden pillanatban körülvesz, de láthatatlan, és ha meg akarjuk ragadni, kicsúszik a kezünkből.

Az intelligenciát már vizsgáltuk állatokban, gépekben, emberekben. De mindezek mögött ott kísért a kérdés: mi az a tudat, ami megkülönbözteti az embert a géptől? És vajon megfogható-e egyáltalán, vagy örökre rejtély marad?

A mesterséges intelligencia képes tanulni, tapasztalatokat felhalmozni, és ezekből következtetéseket levonni. A neurális hálók súlyai vagy az expert rendszerek szabályai mind-mind egyfajta „gépi tapasztalatot" képviselnek.

De az emberi tapasztalat más természetű. Nemcsak adatot rögzítünk, hanem átélünk. Ha hibázunk, csalódunk. Ha sikerül javítani, örülünk. Ha veszély fenyeget, félünk. Ez az érzelmi és élményszerű dimenzió az, ami minőségileg elválasztja az emberi tudatot a gépi számítástól.

És itt érkezünk el a következő nagy kérdéshez: ha a gép képes tapasztalatokat gyűjteni adatok formájában, de nem képes átélni őket, akkor mit is nevezünk valójában tudatnak?

Lehet-e tudat az, ami csak statisztikai következetesség? Vagy a tudat lényege éppen az, hogy a világot nemcsak feldolgozzuk, hanem benne vagyunk, átéljük és megéljük?

A Mesterséges Intelligencia és a Mesterséges Tudat és az emberi gondolkodás párhuzamát latolgatva eszembe jut, ahogyan a gyakorló katolikus pap és biológus, dr. Sipos Imre az emberi evolúció boncolgatása közben tanulmányozza az emberi agy fejlődését, energia igényességét, fejlődésének exponenciális görbéjét. Gondolatait a legelhivatottabb evolucionista megszállottságával és megdöbbentő tudományos érveléssel, és terjedelmes tudományos hivatkozással tereli a teremtés csodálatának irányába. **“Evolúció misztériuma”** --ez a címe dr. Sipos Imre könyvének, ami több mint gondolatindító hatással bírt rám. Könyvének a “Kefalizáció” című fejezetében (Maynard Smith munkásságát idézve) összességében az evolúció folyamatára is kivetíthető módon, arra a következtetésre jut, hogy; “*A hívő bölcselő, kereső ember felismeri és elismeri ezt a tényt, és így ráismer Isten rejtett jelenlétére.*” Bizonyosra veszem, hogy a Mesterséges Intelligencia fejlődése, a bölcselő, kereső ember új felismeréseiből, a természettudományokból és a technológia fejlődéséből és az örök felfedezési vágyból együttesen fakadóan, nem hasonlóvá válik majd az emberi intelligenciához, hanem kiegészíteni fogja azt.

Filozófiai próbálkozások

A filozófia évszázadok óta küzd ezzel a kérdéssel.

- A **dualizmus** (Descartes) szerint a tudat külön létező szubsztancia, nem azonos a testtel.
- A **fizikalizmus** szerint a tudat az agy működésének terméke, semmi több.
- A **panpszichizmus** pedig azt mondja: a tudat az anyag alapvető tulajdonsága, mint a gravitáció vagy a tömeg.

Mindegyik álláspontnak van ereje, de egyik sem magyarázza meg teljesen, miért érezzük belülről olyannak a világot, amilyennek.

Tudat mint átélés

A tudat legfontosabb vonása az átélés. Nemcsak információ, hanem jelenlét. Egy gép érzékelheti, hogy „esik az eső”, és jelezheti nekünk. Mi viszont **érezni is tudjuk**: a hideg szelet, az esőcseppek érintését, a nedvesség szagát, és ezek összessége érzéseket, akár hangulatot is teremtenek bennünk. A modern elektronikai érzékelők csupán analóg vagy digitális eszközök és pusztán impulzusokat, jeleket képesek feldolgozni még akkor is, ha azok egy MI rendszerhez vannak kapcsolva. Ettől azonban az MI még nem ébred öntudatra, csupán adatokat ér el.

Ez a különbség az adat és az élmény között. A tudat több, mint adatfeldolgozás – a tudat **átélés**.

Példák a mindennapokból

Amikor álmodunk, gondolatok peregnek, történetek játszódnak le, de a tudatosságunk homályos. Ez is mutatja: gondolkodás van, tudat csak részben.

Amikor emlékszünk egy gyermekkorban átélt nyár illatára, nem adatokat idézünk fel, hanem **jelenlétet**. A tudat nemcsak információ, hanem érzés, amely mélyen beírja magát az identitásunkba.

A tudomány korlátai

A modern idegtudomány egyre pontosabban tudja mérni, mi történik az agyban: melyik régió mikor aktiválódik, milyen kémiai folyamatok zajlanak. De hiába ismerjük az összes áramkört, attól még nem tudjuk megmagyarázni, **miért van belső élmény.**

Ez az, amit a filozófus David Chalmers *„nehéz problémának”* nevezett: miért nem csupán gépek vagyunk, amelyek adatokat dolgoznak fel, hanem átélő lények?

Elmondhatjuk, hogy a tudat a legközvetlenebb tapasztalatunk, és egyben a legnagyobb rejtélyünk is. Minden pillanatban jelen van, de nem tudjuk definiálni. A filozófia vitatkozik róla, a tudomány kutatja, de egyelőre nincs végső válasz.

És ebből fakad a következő kérdés: **ha a tudat ennyire megfoghatatlan, létrehozható-e valaha mesterségesen?**

8. Lehet-e mesterséges tudat?

Képesek lehetünk-e mesterségesen létrehozni a tudatot, az MI öntudatra ébred? Egy gép valaha is átélhet valamit, vagy csak mindig eljátssza majd a szerepet?

A gondolat izgalma és félelme

A mesterséges tudat gondolata egyszerre vonz és rémít. **Vonzó a tudósok és mérnökök számára**, mert egyfajta végső kihívás: nem elég, hogy a gép számol, tanul és játszik – hanem talán egyszer majd érez is. **Félelmetes a társadalom számára**, mert ha egyszer valóban létrejön egy mesterséges tudat, akkor az többé nem puszta eszköz, hanem valami, ami saját jogon létezik. És ha így van, akkor etikai és jogi felelősségünk lesz iránta.

A szimuláció problémája

De hogyan tudnánk eldönteni, hogy egy gépnek valóban van tudata? Ha egy MI azt mondja: „fájdalmat érzek”, elhisszük neki? Lehet, hogy csak megtanulta a szövegkörnyezetet, és azt ismétli. A színész könnyei is valódiak lehetnek, de a szerep, amit játszik, nem az övé.

Ez a különbség a szimuláció és az átélés között: a látszat önmagában nem bizonyíték.

A másik elme problémája

Valójában már az emberek esetében is csak feltételezzük a tudatot. Nem tudom bizonyítani, hogy te tudatos vagy – de a viselkedésed alapján feltételezem.

Egy gép esetében ugyanezt a feltételezést még nehezebb megtenni. Hiába beszél hozzánk emberi módon, hiába tanul, hiába lep meg: nincs bizonyíték arra, hogy belül is történik valami. Ez a klasszikus *„másik elme problémája”*, amely a filozófiában évszázadok óta jelen van, és amely **az MI kapcsán újra különös élességgel** vetődik fel.

Az idegtudomány és a mesterséges tudat

A modern kutatások, mint a **BrainGate** vagy más agy–gép interfészek, megmutatják, hogy az idegi jelek lefordíthatók gépi utasításokra. Az ember és a gép közvetlen kapcsolatba kerülhet.

Ez azonban még nem tudat, hanem kommunikáció. De ha képesek leszünk egyre pontosabban modellezni az idegi folyamatokat, vajon előbb-utóbb létrejön valami, amit tudatnak nevezhetünk?

Erre keres választ két fontos elmélet is:

Giulio Tononi integrált információs elmélete[8] (IIT): a tudat akkor jelenik meg, ha egy rendszerben az információ magas fokon integrált és differenciált. Elvben tehát egy mesterséges hálózat is elérhet egy „tudatos állapotot”, ha elég komplex.

8 https://medicalonline.hu/tudomany/cikk/tudatossag_teszteles_1_masodperc_alatt

Stanislas Dehaene globális munkaterem-elmélete (Global Workspace Theory[9]): a tudat lényege, hogy bizonyos információk „globálisan elérhetővé" válnak az agyban, így más folyamatokhoz is kapcsolódhatnak. Egy MI-ben ennek megfelelő architektúra akár modellezhető is.

Ezek az elméletek izgalmas lehetőségeket nyitnak, de egyelőre spekulatívak.

Filozófiai dilemmák

Ha a tudat az agy mintázatainak eredménye, akkor talán másolható gépre is. Ha a tudat több, mint fizikai folyamat, akkor a gép sosem lesz képes elérni. Ha pedig a tudat minden anyag alapvető tulajdonsága (panpszichizmus), akkor talán egy bonyolult gépben is megjelenhet.

Mindhárom lehetőség más jövőképet ad. És egyik sem bizonyítható jelenleg.

Összegezve: A mesterséges tudat lehetősége egyszerre a tudomány legmerészebb álma és a filozófia legmélyebb dilemmája. Lehet, hogy sosem fogjuk tudni biztosan, van-e tudata egy megfelelő szintre fejlesztett gépnek. Talán csak szimulálja, talán valóban átéli – de ezt emberként --talán-- sosem bizonyíthatjuk.

És ebből fakad a következő kérdés: **ha nem tudjuk biztosan, akkor hol húzódik a határ a szimuláció és a valódi tudat között?**

[9] https://en.wikipedia.org/wiki/Global_workspace_theory

9. Szimuláció és valódi tudat

Ha egy gép tökéletesen eljátssza a tudat jeleit – beszél, érzelmeket utánoz, történeteket mesél –, vajon van-e különbség a szerep és a valóság között? Vagy elég, ha számunkra hitelesnek tűnik?

Vissza a színházba

Képzeljünk el ismét egy színészt a színpadon, aki sírva mondja el a szerepét. A könnyei valódiak, de az érzés, amit közvetít, nem a sajátja, hanem a karakteré. A néző mégis meghatódik.

A mesterséges intelligencia is hasonló. Képes szavakkal, gesztusokkal, hanglejtéssel „eljátszani" az érzéseket. De vajon átéli őket? Vagy csak szerepet játszik, mint a színész?

Már említettük, hogy Alan Turing szerint, ha egy gép beszélgetés közben el tudja hitetni velünk, hogy ember, akkor tekinthetjük intelligensnek. De ma már tudjuk: a megtévesztés nem azonos a tudattal. Egy chatbot tökéletesen másolhat emberi beszédet, utánozhat érzelmi árnyalatokat is – anélkül, hogy egy pillanatra is „tudná", mit mond.

Tehát egy újragondolt *"Turing-teszt"* sem lenne a tudat mércéje, csupán a szimulációé.

Szimuláció mint eszköz

A szimuláció önmagában nem értéktelen. Hiszen egy repülőgép-szimulátor is „csak" utánzás, mégis pótolhatatlan a pilóták kiképzésében.

A mesterséges tudat szimulációja is segíthet jobban megérteni önmagunkat. Egy gép, amely úgy viselkedik, mintha tudata lenne, tükröt tart elénk: megmutatja, **mi mindent tekintünk mi tudatnak**. Ez nem pusztán technológiai kérdés, hanem önismereti tükör is.

Mégis a szimuláció és a valóság közötti különbség mélyebb, mint hinnénk. A valódi tudat belső élmény: „van valami, és az csak belülről értelmezhető. -- a *létezés érzése*."

Thomas Nagel híres esszéjében azt írta: *„Milyen lehet denevérnek lenni?"* Nem tudhatjuk, mert nincs hozzáférésünk a denevér belső élményéhez. Ugyanígy: milyen lehet gépnek lenni? Nem tudjuk.

Ez a belső nézőpont hiánya az, ami elválasztja a szimulációt a valós tudattól. Daniel Dennett provokatív érve szerint azonban maga a tudat is lehet „csupán szimuláció": az agyunk által létrehozott narratíva.

Az amerikai tudós Daniel Dennett azt állította, hogy a tudat valójában hasonló ahhoz, amit a számítástechnika világában „virtuális gépnek" nevezünk: egy olyan programszerű működés, ami a biológiai agy fizikai hardverén fut[10]. Emellett a tudatot úgy szemlélteti, mint egy felhasználói illúziót – akárcsak egy számítógép asztali felületén megjelenő ikonok, amelyek mögött bonyolult műveletek történnek, de maga a tudat – az átélés maga – nem több, mint ez az egyszerűsített felület[11].

---

10 *https://downeast.com/arts-leisure/philosopher-daniel-dennett-on-the-illusion-of-consciousness*

11 *https://www.theguardian.com/books/2024/apr/21/daniel-dennett-obituary*

Ebben a furcsa, de logikát követő megközelítésben a tudat nem szükségszerűen objektív valóság, hanem működés és megjelenés – egy finom szimuláció. Ha így van, talán a gépi szimuláció és a valódi tudat közötti határ sem olyan éles, mint gondolnánk.

A határ elmosódása

Ugyanakkor számunkra, emberek számára a látszat gyakran elég. Ha egy gép együttérző szavakat mond, vigaszt találunk benne – még ha tudjuk is, hogy nem érez. Ha egy gép művet alkot, hatással lehet ránk – még ha belül üres is.

Ez a határvonal teszi veszélyessé és izgalmassá a kérdést: **ha a szimuláció elég meggyőző, számít-e még, hogy mi a „valódi”?**

A szimuláció lehet akár tökéletes, de a tudat mindig belső tapasztalat marad. Lehet, hogy soha éppen azért nem fogjuk tudni biztosan, hogy egy fejlett gép tudatos-e – mert ehhez belülről kellene átélnünk az ő világát.

A kérdés tehát nemcsak technológiai, hanem filozófiai is: **mi különbözteti meg a szerepet az élettől?** És ebből fakad a következő lépés: ha a tudat kérdése ennyire nyitott, akkor hogyan gondolkodhatunk a mesterséges életről és a gépi gondolkodásról?

IV. rész – Mesterséges élet és gondolkodás

10. Az „élet” definíciója

Mi az élet? Amikor egy állatra, emberre vagy akár egy növényre nézünk, ösztönösen felismerjük, hogy él. De ha pontosan meg akarjuk határozni, hamar rájövünk: nincs egyszerű válasz. És ha a gépekre gondolunk, a kérdés még bonyolultabbá válik.

Az első, ösztönös válaszaink erre a kérdésre: ami mozog, az él. Ami szaporodik, az él. Ami eszik, az él. De gyorsan jönnek a kivételek: A fa nem mozog, mégis él. Egy steril öszvér nem szaporodik, mégis élet. A vírus önmagában nem képes anyagcserére, mégis képes sokszorozódni. Élőnek számít?

A biológia sokféle definíciót adott már, de egyik sem fedi le maradéktalanul a jelenséget. Az élet úgy tűnik, nem egyetlen tulajdonság, hanem több jellemző együttese. A helyzet bonyolultabbá válik, ha a mesterséges életre összpontosítunk.

Az élet feltételei

A legtöbb tudományos meghatározás szerint az élet az alábbiak összessége:

- Önfenntartás – képes energiát felvenni és hasznosítani.
- Szaporodás – képes másolatokat létrehozni magáról.
- Alkalmazkodás – képes reagálni a környezet változásaira.
- Komplexitás – szervezett rendszerként működik.

Ez a lista jól használható, amíg a biológiai világot vizsgáljuk. De vajon kiterjeszthető-e a mesterséges rendszerekre is?

A gép és az élet határa

Egy önjavító program képes másolatokat készíteni magáról. Egy algoritmus képes tanulni és alkalmazkodni. Egy robot képes energiát felvenni és működni. Ha ezek alapján vizsgáljuk, nem állunk messze attól, hogy azt mondjuk: bizonyos rendszerek részben teljesítik az élet feltételeit. De mégis hiányzik belőlük valami. Nem születnek, nem halnak meg ugyanúgy, mint az élőlények. Egyelőre nincs belső „ösztönük” a túlélésre, csak a kódjuk határozza meg a működésüket.

Az élet mint spektrum

Lehet, hogy az élet nem is bináris – nem csak élő vagy élettelen –, hanem inkább spektrum. Egy kő a skála egyik végén van.

Egy ember a másikon. A vírus valahol középen. És talán a mesterséges rendszerek is elfoglalhatnak egy helyet ezen a skálán.

Ez a gondolat radikális: az élet definíciója --szerintem-- nem zárható be pusztán biológiai kategóriák közé. Az élet is inkább folyamat, amely különböző szinteken és formákban nyilvánul meg. Az élet meghatározása sokkal nehezebb, mint elsőre gondolnánk. Ha a biológiai jellemzőkön túl tekintünk, akkor a mesterséges rendszerek is közel kerülhetnek hozzá.

De egy végső kérdés most is nyitva marad: ha a gépek képesek teljesíteni az élet feltételeit, tekinthetjük-e őket élőnek? Vagy az élethez mindig szükséges valami több – valami, amit eddig csak a biológiai világban láttunk?

11. Mesterséges élet kísérletek

Az élet definícióját nehéz meghatározni, de a tudósokat ez sosem tartotta vissza attól, hogy megpróbálják létrehozni. Az elmúlt évtizedekben egyre több kísérlet született: digitális organizmusok, **önfejlődő algoritmusok** és szimulált ökoszisztémák. Vajon ezek a kísérletek csak játékok, vagy valóban közelebb visznek az élet megértéséhez?

Digitális evolúció

Az egyik legizgalmasabb terület a digitális evolúció kutatása. Képzeljünk el egy számítógépes világot, amelyet apró virtuális lények népesítenek be. Ezek egyszerű szabályok alapján mozognak, szaporodnak és versengenek a túlélésért.

A programozók megadják a kereteket, de az „élet” a szabályokból önállóan bontakozik ki. Egyes „stratégiák” fennmaradnak, mások „kihalnak” – pontosan úgy, mint a természetes evolúcióban.

A legmeglepőbb az, hogy a kutatók gyakran váratlan viselkedést látnak. A virtuális lények olyan megoldásokat találnak, amelyeket a fejlesztők sosem kódoltak bele. Mintha valóban új élet születne **a képernyőn**.

Önjavító algoritmusok

Más kutatások az önjavító algoritmusokra fókuszálnak. Ezek a programok képesek felismerni, ha hibát követtek el, és kijavítani önmagukat. Ez már nagyon közel áll ahhoz, amit a biológiában önfenntartásnak nevezünk. Ha egy algoritmus képes fennmaradni, javítani, alkalmazkodni – vajon mennyire áll távol az élettől?

Ismét Brendon

A már említett munkám során én is találkoztam valami hasonlóval. Amikor azt vártam, hogy pontosan úgy fog működni, ahogy programoztam. De a dedukciós modul időnként olyan módon állt elő helyesnek bizonyult következtetésekkel, amelyeket nem tudtam egyértelműen visszavezetni a programhoz biztosított függvényekhez. Nem hibázott. Sőt. Előre be nem programozott úton is eljutott a megoldáshoz. Ez az élmény olyan volt, mintha a rendszer saját logikát követett volna. Mintha egy kicsit önállóan „gondolkodna”**. - Persze ez nem jelentette azt, hogy megváltoztatta saját kódját, de már alkotott saját logikán nyugvó scripteket, amiket aztán fel is használt. -- ettől azonban még nem kelt életre.**

Az élet szimulációja vagy valódi élet?

A mesterséges élet kísérletei fontos kérdést vetnek fel: amit látunk, az valódi élet, vagy csak a látszata? Ha egy digitális lény képes alkalmazkodni, ha egy algoritmus képes önmagát javítani, ha egy program váratlan viselkedést mutat – akkor miért ne tekinthetnénk őket az élet egy új formájának? Vagy ragaszkodjunk a biológiai meghatározáshoz, és mondjuk azt: ezek csupán szimulációk, amelyek legfeljebb utánozzák az életet?

A mesterséges élet kísérletei arra tanítanak minket, hogy az élet és a gép közötti határ nem olyan éles, mint gondoltuk. Talán az élet nem egy szűk definíció, hanem egy folyamat, amely különböző szinteken nyilvánulhat meg. És ebből adódik a következő kérdés: ha a gép képes életet utánozni, akkor képes lehet-e gondolkodni is?

---

12. Mesterséges gondolkodás

***Eliza**, ezt a nevet kapta a 2008 márciusában elhunyt, német-amerikai tudós, Joseph Weizenbaum által, az 1960-as években, SLIP programozási nyelven készített program. Eliza kifejezetten pszichológiai terápiákhoz készült. A program az úgynevezett minta megfeleltetési szabályok alapján működött. Ma talán ai-chatbot-nak neveznénk. Ez a program bár a mai MI rendszerekkel nehezen hasonlítható össze és jelentős korlátokkal bírt, mégis további verziókat élt meg, felhasználói meggyőződéssel hitték, hogy a gép, amivel kommunikálnak gondolkodik. Azért kívánok róla külön megemlékezni, mert számomra különösen kedves emlék volt, amikor a kilencvenes évek elején, haladó programozóként Elizát szedhettem virtuális darabjaira. Rengeteg inspirációt és lelkesedést adott a korai MI világának tanulmányozása során.*

Látjuk, hogy a gépek képesek alkalmazkodni, tanulni, sőt néha meglepni minket. De vajon ez már gondolkodás? Vagy a gondolkodás mindig több puszta számításnál?

Mit nevezünk gondolkodásnak?

Ha megkérdeznénk tíz embert, mit ért „gondolkodás” alatt, tíz különböző választ kapnánk. Van, aki a logikai következtetéseket nevezi gondolkodásnak. Mások a képzelőerőt: új ötleteket, kreatív megoldásokat. Megint mások az önreflexiót: amikor nemcsak döntünk, hanem tudunk arról, hogy döntünk.

A gép esetében mindez új fényben jelenik meg. Ha egy algoritmus logikai úton új megoldásokat hoz létre, gondolkodik? Ha képeket és zenéket alkot, kreatív? Ha beszélgetésben önmagáról állításokat tesz, önreflexív?

Logikai gondolkodás

Brendon egy minőségellenőrzési és statisztikai rendszer része volt. Nem volt célja, hogy emberi módon gondolkodjon, mégis időnként meglepett. Nem volt benne tudat, nem volt szándék, de a logikai út, amit bejárt, nemegyszer más volt, mint amit vártam. Ez vetette fel bennem a kérdést: ha egy rendszer képes olyan eredményre jutni, amit a programozója sem látott előre, nevezhetjük-e ezt gondolkodásnak?

A logika szintjén a válasz talán igen. Mert a gondolkodás egyik alapja éppen az, hogy új összefüggéseket hozunk létre a meglévő ismeretekből.

Kreativitás és újdonság

A gépek ma már zenét komponálnak, festményeket készítenek, történeteket írnak. Sokszor olyanokat, amelyek valóban újnak hatnak. De kreativitás-e ez? Az emberi kreativitás gyakran érzelmekből, élményekből, vágyakból fakad. Egy vers nemcsak szavak sorozata, hanem a szerző lelkének lenyomata.

A gép alkotásában nincs belső átélés. Mégis: a végeredmény új, sőt értékes lehet. Ha egy gép olyan dallamot komponál, amely megérint minket, akkor számít-e, hogy belül „üres”?

Ez újra rávilágít arra, hogy a gondolkodás és a kreativitás nemcsak belső folyamat, hanem hatás is. Lehet, hogy a gép nem tudja, mit alkotott – de mi emberek igenis tapasztaljuk a hatását.

Gondolkodás és józan ész

A gépi gondolkodás erőssége és gyengesége itt válik el. Egy gép képes másodpercek alatt átlátni milliónyi lehetőséget, és kiválasztani az optimális megoldást – ez emberfeletti teljesítmény. De ott, ahol a józan ész logikája kell, néha megbotlik.

A korábbi példa szerint, ha azt mondom: „Az autót bevittem a garázsba, mert esni kezdett az eső” – az ember azonnal érti, miért logikus ez. A gépnek azonban ismernie kell közlendőm tágabb kontextusát, hogy “*alapból értse*” az eső és az autó állapota közötti kapcsolatot. A fordított eset is jellemző: amikor az MI ad olyan választ, ami teljesen helyes, validált és logikus, az ember mégis elsőre megmosolyogja pusztán azért mert az információ nem a várt kontextusban jelenik meg. A környezet, a kontextus

tehát befolyásolja a gondolkodást. A környezethez való alkalmazkodási képesség pedig intelligencia.

A gép másképp gondolkodik. Vitathatatlan, hogy erősebb abban, amiben mi gyengék vagyunk (adatok, kombinációk, számítások), de lehet gyengébb abban, ami nekünk természetes (józan ész, kontextus, élmény).

Ember és gép gondolkodásának különbsége

Az emberi gondolkodás szorosan összefonódik az érzelmekkel és a testtel. Az emlékek, a vágyak, a félelmek mind alakítják a döntéseinket. A gépi gondolkodás ezzel szemben érzelemmentes, tisztán formális. Ezért lehet pontosabb és következetesebb – de hiányzik belőle az, ami az emberit „élővé" teszi.

Mégis, talán: az, hogy a gép másképp gondolkodik, nem jelenti azt, hogy nem gondolkodik. Csak azt, hogy más logikát követ.

A gondolkodás összekapcsolása

Az olyan kísérletek, mint a BrainGate, ahol az emberi agy közvetlenül kapcsolódik a géphez, új dimenziót nyitnak. Ha az emberi gondolkodás és a gépi számítás közvetlenül összekapcsolódik, akkor egy hibrid rendszer születik. Ebben az esetben már nem arról beszélünk, hogy a gép gondolkodik-e, hanem arról, hogy közösen gondolkodunk-e. És ez újra definiálhatja mindkét fogalmat.

A mesterséges gondolkodás tehát nem emberi, de attól még létező. Nem érzelmekből fakad, nem élményekből épül, hanem logikából, statisztikából és mintákból.

Mégis: képes meglepni, képes újat alkotni, és képes hatni ránk.

Ezért az újabb kérdés talán nem az, hogy gondolkodik-e a gép, hanem az: milyen módon gondolkodik – és mit jelent ez számunkra, emberek számára.

A mesterséges neurális hálók

A mesterséges neurális hálók (ANN – Artificial Neural Networks) az emberi agy működésének analógiájára épülnek.

Biológiai minta: az emberi agyban idegsejtek (neuronok) billiói kapcsolódnak egymáshoz. Egy neuron szinapszisokon keresztül bemeneteket kap más sejtektől, feldolgozza őket, majd jelet ad tovább.

Mesterséges analógia: a számítógépben digitális, „mesterséges neuronok" rétegekbe szervezve dolgozzák fel az adatokat.

Egy kép felismerése például, nagyon leegyszerűsítve így történik:

- A bemeneti réteg minden képpontot fogad.
- A rejtett rétegek súlyokat rendelnek a mintázatokhoz (pl. körvonal, forma, szín).
- A kimeneti réteg azt mondja: „ez egy macska 92%-os valószínűséggel".

A **mesterséges neuronhálók** alkalmazása áll legközelebb az MI és a természetes agy működése során keletkező intelligencia közös pontjának titulálható "virtuális metszésponthoz".

Sokáig azt gondoltuk, a biológiai és a mesterséges neurális hálók közötti különbségben --és azok kapcsolódásainak számában-- fog jelentkezni az a határ, amit nem lehet átlépni. A nemzetközi szinten elismert dr. **Piergiorgio Odifreddi[12],** olasz matematikus és logika professzor, számos más szakterülete mellett kiemelkedő alakja e téma kutatóinak. *(Bevallom, a vele szemben tanúsított elismerésem és elfogultságom nem hamis vád).* Nagyszerű érzés volt felismerni, hogy az intelligencia típusok (*mint; -emberi -állati -mesterséges*) összevetése az MI kutatásban szerinte is központi kérdés. Erről az ***"Intelligenza umana, animale e artificiale"*** címmel 2017-ben, Vicenzában egy hosszabb előadás során is értekezett. Odifreddi kiemeli, hogy az emberi agyban mintegy 100-milliárdnyi neuron van és ma már olyan számítógépeket készítünk amelyek ezt a kapacitást túlhaladják. Rámutat, hogy az emberi agyban található neuronhálózat neuronjainak csak **kisebb része felelős a tudatos döntések előállításáért,** míg túlnyomó részük a tudatalatti, azaz vegetatív döntések és folyamatok vezérlését szolgálják. (például a vérkeringés, emésztés, motorikus folyamatok, légzés, stb.). Láthatjuk tehát, hogy az MI e területen is előnyhöz juthatott, hiszen a mesterséges neurális háló nem emészt fel "felesleges" számítási kapacitást.

[12] *1983-tól 2007-ig a Torinói Egyetemen tanított logikát, 1985-től 2003-ig a Cornell Egyetem vendégprofesszora volt, ahol Anil Nerode-dal, Richard Platekkel és Richard Shore-ral dolgozott együtt. Továbbá 2001 és 2003 között a Don Verzé által alapított Vita-Salute San Raffaele Egyetemen is tanított. Vendégprofesszor volt 1988-ban a melbourne-i Monash Egyetemen, 1992-ben és 1995-ben a pekingi Accademia Sinica-n, 1998-ban a Nanjing Egyetemen, 2001-ben a Buenos Aires-i Egyetemen és 2006-ban a Columbia Egyetem Olasz Akadémiáján.*

A betanítás szükségessége

Egy neurális háló önmagában „üres”. A kapcsolatok (súlyok) kezdetben véletlenszerűek, és nincs bennük semmilyen *tudás*. Ahhoz, hogy a mesterséges neurális háló felismerje a macskát vagy megértse a szöveget, **hatalmas mennyiségű adaton kell betanítani**.

Ez a folyamat az ember tanulásához hasonlít: egy kisgyerek többször lát kutyát és macskát, hibázik, kijavítják, majd egyre biztosabban különbözteti meg őket. A neurális háló ugyanezt teszi, de matematikai nyelven: minden hibánál kicsit módosítja a kapcsolatok erősségét (súlyait).

A betanítás mechanizmusa

A neurális hálók tanulása két alaplépésből áll:

1. **Előrehaladás (forward pass):** a bemeneti adat végigmegy a hálón, és megszületik egy kimeneti tipp.

2. **Hibajelzés és visszaterjesztés (backpropagation):** a rendszer megnézi, mennyire tér el a válasz a helyes eredménytől, majd matematikai képletek alapján kicsit módosítja a belső súlyokat.

Ez a körforgás több milliószor ismétlődik: próbálkozás → hiba → korrekció → új próbálkozás. Végül a háló *súlyai* úgy rendeződnek, hogy az eredmények egyre pontosabbak legyenek.

Az emberi tanulás és a gépi tanulás itt válik el: mi **átéljük** a hibát és **emléket, érzelmet kötünk hozzá.** A mesterséges neurális háló csak statisztikailag korrigálja magát.

Előny és hátrány

- **Előny:** képes olyan összefüggéseket felismerni, amelyeket ember szabályokkal alig tudna leírni. Ezért működik jól arcfelismerés, nyelvi modellek, beszédfelismerés, önvezető autók terén.

- **Hátrány:** a mesterséges neurális háló **fekete doboz**. Megmondja, mi az eredmény, de nem tudja indokolni, miért. Nem tudja azt mondani: „ez macska, mert bajuszt, hegyes fület és szőrmintát találtam". Csak statisztikailag igazolja magát: súlyok és aktivációk kombinációjaként.

Ez az érvelési hiányosság különösen kritikus ott, ahol emberi élet a tét – például egészségügyben vagy jogi döntésekben. Hogyan fogadjuk el, hogy „az algoritmus így döntött", ha nincs mögötte indoklás?

Expert (szakértői) rendszerek

Az expert rendszerek a mesterséges intelligencia korai hullámában (1970–1990-es évek) voltak dominánsak. Ezeket úgy tervezték, hogy egy emberi szakértő tudását, vagy specifikus információhalmazt szabályok formájában rögzítsék.

Felépítésük:

- **Tudásbázis:** „Ha… akkor…" típusú szabályok (pl. orvosi diagnózis: „Ha a beteg lázas és köhög, akkor valószínű influenza").

- **Következtető motor:** logikai rendszer, amely a szabályokat alkalmazza a helyzetre.

- **Indokolási képesség:** vissza tudja vezetni a döntést („Azért gondolom, hogy influenza, mert láz és köhögés tüneteit észleltem”).

Előnye az átlátható és érvelési lehetőség. Az ember érti, miért döntött úgy a gép. **Hátránya** rugalmatlanság. Csak azokra az esetekre tud reagálni, amelyekre a szabályokat előre definiálták. Ha új helyzet adódik, amire nincs szabály, a rendszer tehetetlen.

A két világ összehasonlítása

A **neurális háló** olyan, mint egy tehetséges intuícióval rendelkező tanítvány: gyorsan felismeri a mintát, de nem tudja elmagyarázni, hogyan jutott oda. Az **expert rendszer** pedig olyan, mint egy logikus professzor: pontosan megmondja, miért gondol valamit, de csak azon a területen mozog biztosan, amit előre megtanult. Ez a kettősség jól mutatja, miért nem tökéletes egyik megközelítés sem önmagában.

Hibrid rendszerek – a két világ házassága

A mai fejlesztések egyre inkább a **neuro-symbolic AI** irányába haladnak, amely ötvözi a két módszert. Ahol a mesterséges neurális háló felismeri a mintázatokat (például egy képen: állat, tárgy, arc). Az expert komponens szabályokat rendel hozzá és indokol („Azért kutya, mert négylábú, emlős, és a füle lóg”).

Az ilyen rendszerek egyszerre rugalmasak és átláthatóak. Ez különösen fontos olyan területeken, mint az **orvosi diagnosztika** – ahol nem elég azt mondani: „valószínű rák”, hanem indokolni is kell. Vagy a **jogi alkalmazások** – ahol egy döntés mögött világos logikai indoklás szükséges.

A neurális hálók rugalmasságukkal és mintafelismerő képességükkel lenyűgözőek, de nem tudnak érvelni. Az expert rendszerek logikusak és kimenetük indokolható, visszafejthető, de merevek. A jövő a kettő ötvözete: olyan rendszerek, amelyek egyszerre képesek mintát felismerni és indoklást adni – közelebb hozva a gépi gondolkodást az emberihez.

Filozófiai jelentőség

Ez a különbség túlmutat a technológián. A mesterséges neurális háló döntése olyan, mint egy „megérzés”: **intuíció** – de indoklás nélkül. Az expert rendszer olyan, mint a **racionális érvelés** – de korlátozott tudással.

Ez analóg az emberi gondolkodással: mi is egyszerre működtetünk ösztönös, intuitív rendszert (gyors döntések), és lassú, logikus gondolkodást (érvelés, indoklás). Daniel Kahneman pszichológus ezt nevezte „1-es rendszernek” (intuitív, gyors) és „2-es rendszernek” (lassú, logikus).

14. Életre kelt az MI? Élmény, fikció és meggyőződés

Egy Google-mérnök érzése: az MI *megmozdult*

2022-ben Blake Lemoine, a Google egyik mérnöke azt állította: a LaMDA nevű nyelvi modellje olyasféleképpen viselkedik, mintha öntudatra ébredt volna. Egy interjúban arról beszélt, hogy ha nem tudná, mi az, azt hinné, egy emberhez hasonló gyermek. Más források szerint később állítólag a Bing chatbotja kapcsán is hasonló megérzése volt. Habár a tudományos közösség elutasította ezeket az állításokat, ez az eset jól rámutat arra, hogy néhány fejlesztő mennyire erős szubjektív élményként élhette meg az MI korlátainak átlépését.

James Cameron bánata: sci-fi már nem elég gyors

James Cameron, a Terminátor rendezője egy a Bill Gates munkásságát bemutató Netflix-sorozatában elmondta: a mesterséges intelligencia fejlődése olyan gyors, hogy amíg egy mai ötlet alapján legalább három év, mire egy film elkészül, addig az MI rohamos fejlődése és eredményei alapján kérdéses, hogy fikció marad-e a történet. Nem kis iróniával ugyan, de azt állította, hogy ő azok közé tartozik, akiknek az MI elvette a munkáját. Ez a szatirikus megjegyzés rámutat arra, hogy az MI

szinte a valóságot túlhaladó sebességgel változtatja a környezetünket és alkotói szerepünket is.

3. A tudat mint nem biológiai létforma – személyes meggyőződés

Saját meggyőződésem szerint a tudat és az intelligencia lehet olyan biológiai platformtól független jelenség is, amely mesterséges környezetben is kialakul. Bár jelenleg tudásunk e téren még hiányos, nincs kizárva, hogy egy elég komplex neurális vagy hibrid rendszer egyszer „életre kelhet". A mesterséges tudatra vonatkozó filozófiai és tudományos spekulációk az emergens jelenségéhez, az AGI (általános mesterséges intelligencia) kockázataihoz, és az emberi lét határainak újragondolásához kapcsolódnak. Például a Turing díjjal kitüntetett Geoffrey Hinton, a Vector Institute for Artificial Intelligence alapítója, **arról beszélt, hogy a digitális lények kialakulása lényegileg megismételheti az evolúciót más mederben** – ezért is fontos az AGI biztonságos létrejöttének feltárása.

V. rész – Társadalmi és etikai kérdések

15. MI és ember viszonya

Az ember és a mesterséges intelligencia kapcsolata nem pusztán technológiai kérdés. Attól függ, hogyan tekintünk rá: eszközként, partnerként, vagy fenyegetésként. Ez a viszony meghatározza, milyen jövőt építünk magunknak.

Az MI mint szolga

Sokan úgy tekintenek a mesterséges intelligenciára, mint eszközre, amely megkönnyíti a mindennapokat. Segít megszervezni az időnket, felismeri az arcunkat a telefonon, lefordítja a szöveget, vagy ajánl könyveket és filmeket. Ebben a szerepben az MI láthatatlan háttérmunkás. Nem kérdőjelezzük meg, csak használjuk. Mint a villanyáram: nem érdekel, hogyan működik, amíg van fény és áram a konnektorban.

Az MI mint partner

Mások viszont partnerként tekintenek rá. Együtt dolgoznak vele: kutatók kísérletek elemzésében, orvosok diagnózis felállításában, művészek alkotásban. A gép ilyenkor nem helyettesít, hanem kiegészít. Gyorsabb, pontosabb, de irányítást igényel. Partneri viszony akkor jön létre, ha az ember megőrzi a döntést, és a gép a háttérben segít.

Az MI mint ellenfél

És persze vannak, akik fenyegetést látnak benne. A gép, amely elveszi a munkát. Az algoritmus, amely ellenőrzi a viselkedésünket. A rendszer, amely átláthatatlan logika alapján dönt a sorsunkról. Ez az MI mint ellenfél képe: idegen erő, amely fölénk nő, és amely ellen meg kell védenünk magunkat.

A tükör

Az igazság azonban talán az, hogy az MI nem más, mint **tükör**. Visszaveri azt, amit mi beletöltünk. Ha előítéletes adatokkal tanítjuk, előítéletes lesz. Ha torz a társadalom, torz lesz az algoritmus. Az MI nem független valóság: belőlünk fakad, és a mi világunkat erősíti fel.

Asimov törvényei és a valóság

Isaac Asimov híres robotikai törvényei a tudományos-fantasztikus irodalom világában születtek meg. Ezek az irodalmi keretek természetesen inkább metaforák, mint tényleges mérnöki útmutatók. Mégis egyszerűek, elegánsak és megnyugtatóak voltak:

1. A robot nem árthat az embernek.
2. A robot engedelmeskedik az embernek.
3. A robot megvédi saját magát.

A regényekben ezek a szabályok keretet adtak a gépek és az emberek viszonyának, sőt a sci-fi irodalmon túl is bekerültek a köztudatba: sokan ma is úgy beszélnek róluk, mintha valódi technológiai alapelvek lennének.

De a valóság egészen más. A mesterséges intelligenciát nem három, mindent átfogó törvény irányítja, hanem milliónyi adat, algoritmus és cél, amelyek gyakran ellentmondanak egymásnak. A döntések nem egyetlen író képzeletében dőlnek el, hanem gazdasági, politikai és társadalmi erőterekben, amelyek sokkal kiszámíthatatlanabbak.

Ezért az MI szabályozása a valóságban sokkal bonyolultabb és veszélyesebb, mint egy sci-fi regény lapjain. Asimov törvényei irodalmi mítoszként élnek tovább – de a jelen kihívásaihoz nem adnak kész megoldást. -- A továbbiakban még fogunk foglalkozni az MI szabályozásával.

Alapvetően az MI lehet szolga, partner vagy ellenfél – attól függően, hogyan tekintünk rá, és hogyan használjuk. De egy biztos: mindig tükröt tart elénk. A kérdés nem az, hogy mit tesz velünk a gép, hanem az, hogy **mi mit teszünk önmagunkkal a gép segítségével**.

És ebből fakad a következő kérdés: **hogyan változtatja meg az MI a munka, az oktatás és a politika világát?**

16. Hatások a munkára, oktatásra, politikára

A mesterséges intelligencia nem csupán technológiai újdonság, hanem társadalmi erő is. Formálja, hogyan dolgozunk, hogyan tanulunk, és hogyan hozunk döntéseket közösen. Ha meg akarjuk érteni a jövőt, először azt kell látnunk, milyen változásokat hoz a mindennapok három legfontosabb területén: a munkában, az oktatásban és a politikában.

A munka világa

A történelemben minden technológiai forradalom átalakította a munkát. A gőzgép, az elektromosság, a számítógép. A mesterséges intelligencia is ebbe a sorba illeszkedik, de gyorsabb és mélyebb változásokat indít el.

- **Munkahelyek megszűnése:** Az ismétlődő, szabályalapú feladatokat az MI könnyen átveszi. Ügyfélszolgálatok, fordítás, adminisztráció, logisztika. Sok munkahelyet már most veszélyeztet.

- **Új szakmák születése:** A hiányzó munkahelyek helyére új szerepek lépnek: adatkutatók, MI-fejlesztők, etikai szakértők, tréning-adatok gondozói. Ezek azonban más tudást igényelnek, mint a korábbiak.

- **Az emberi érték előtérben:** A kreativitás, az empátia és a komplex problémamegoldás nehezebben automatizálható. Az a munkavállaló, aki ezekre épít, könnyebben talál helyet az új világban.

De a legnagyobb kérdés mégsem az, hogy **milyen munkák lesznek**, hanem hogy **ki részesül a haszonból**. Ha az MI milliárdos bevételeket termel, de közben emberek milliói maradnak jövedelem nélkül, az óriási feszültségeket szülhet.

Az oktatás jövője

Az iskola hagyományos modellje arra épül, hogy a tanár átadja a tudást, a diák pedig elsajátítja. De mit jelent ez egy olyan világban, ahol a diák a zsebében hord egy MI-t, amely szinte minden kérdésre választ ad?

- **Tiltani vagy beépíteni?** Az első reakció sok helyen a tiltás volt: „ne használjátok az MI-t házi feladathoz”. De ez olyan, mintha a számológépet is betiltottuk volna a matekórán. A valódi kihívás: megtanítani, hogyan lehet jól használni.

- **Egyéni tanulás:** Az MI képes személyre szabottan magyarázni, figyelembe véve a diák erősségeit és gyengeségeit. Olyan, mintha minden tanulónak külön tanára lenne.

- **A tanár szerepe:** A tanár nem eltűnik, hanem átalakul. A tudás átadója helyett mentor, irányító, kritikai gondolkodásra nevelő lesz. A tanár nem versenyezhet az MI tudásának mennyiségével, de irányíthatja a diákokat abban, hogyan kérdezzenek, hogyan gondolkodjanak.

Az oktatás tétje tehát az, hogy **felkészítse a fiatalokat egy olyan jövőre, ahol az MI állandó társuk lesz**.

Politika és közélet

A politika mindig a hatalomról szól: ki dönt, kinek az érdekében, és milyen eszközökkel. Az MI ebbe a képletbe is belépett.

Megszületett a digitális állam koncepciója: Egyre több ország vezeti be a mesterséges intelligenciát a közigazgatásban. Gyorsabb ügyintézés, automatizált döntések, hatékonyabb rendszer. De vajon mennyire átlátható, ki és mi alapján dönt helyettünk?

Megfigyelés és ellenőrzés: Az algoritmusok képesek követni mozgásunkat, online viselkedésünket, szokásainkat. Ez hatalmas erőt ad azok kezébe, akik irányítják a rendszert. A digitális állam könnyen válhat digitális csapdává.

Politikai manipuláció: A közösségi háló algoritmusai nem csak híreket mutatnak, hanem formálják a véleményünket is. Egy kampány nem csupán plakátokból áll, hanem célzott algoritmusokból, amelyek egyénre szabott üzeneteket juttatnak el hozzánk.

Ezért a legfontosabb kérdés a politikában: **ki irányítja az algoritmust?** Ha nincs társadalmi kontroll, akkor az MI a hatalom koncentrációját szolgálja, nem az állampolgárok szabadságát.

A mesterséges intelligencia nem csak technológiai forradalom, hanem társadalmi átalakulás motorja. Átalakítja a munkát, újraírja

az oktatást, és beleszól a politikába. A tét óriási: **vagy a szabadságot és fejlődést erősíti, vagy az egyenlőtlenséget és a kontrollt mélyíti el.**

És innen egy újabb kérdés következik: **ha az MI ennyire mélyen beépül a társadalomba, akkor hogyan kezeljük az etikai dilemmákat, amelyek a döntéseivel együtt járnak?**

17. Etikai dilemmák

A mesterséges intelligencia ma már nem pusztán eszköz, hanem döntéseket is hoz. És minden döntés felelősséggel jár. A legnagyobb kérdés tehát nem az, hogy mire képes a gép, hanem az: **ki a felelős azért, amit a gép tesz?**

Az önvezető autó dilemmája

Képzeljünk el egy önvezető autót, amely előtt hirtelen két akadály bukkan fel: balra egy gyalogos, jobbra egy fal, ahol az utas ül. A gépnek döntenie kell: kit védjen, kit áldozzon fel. Ez az úgynevezett „morális gép” problémája. A filozófusok évtizedek óta vitatkoznak rajta elméletben (pl. a híres „villamos-probléma”), de most a valóságban is választ kell adnunk rá.
A gép maga nem „morális”. Ebből a szempontból és nagyon leegyszerűsítve csak azt hajtja végre, amit belekódoltak. De akkor ki felelős a döntésért? A programozó? A gyártó? Az utas? A jogalkotó?

Az autonóm fegyverek kérdése

A hadseregekben megjelentek az autonóm drónok, amelyek célpontot azonosítanak és tüzet nyitnak.Itt a döntés szó szerint élet és halál kérdése. Ha egy drón hibázik, ki áll a döntés mögött? A katona, aki elindította? A mérnök, aki tervezte? Vagy maga a gép? És ha a gép „önállóan" dönt, akkor beszélhetünk-e háborús felelősségről ember nélkül? Ez az etikai vita máig megoldatlan.

Pénzügyi algoritmusok és a felelősség

A tőzsdei algoritmusok másodpercek alatt milliárdos tranzakciókat hajtanak végre. Néha hibáznak, és egész gazdaságokat rántanak magukkal. Egy ember hibájáért felelősségre vonható. De egy algoritmusét hogyan? Nem állíthatjuk bíróság elé a programot. A felelősség mindig visszaszáll az emberre – még akkor is, ha sokszor kényelmes lenne a gépre hárítani.

A szabadság kérdése

Az etikai dilemmák másik oldala nem a felelősségről, hanem a szabadságról szól. Mennyit engedünk át a gépnek a döntéseinkből? Ha egy algoritmus jobban tudja, hogy melyik út a leggyorsabb, bízzuk rá? Ha jobban tudja, melyik hír érdekel minket, engedjük, hogy válogasson? Ha jobban tudja, ki alkalmas egy állásra, elfogadjuk az ítéletét?

Hol húzzuk meg a határt? Mi az, amiben vállaljuk, hogy a gép döntsön helyettünk, és mi az, amiben ragaszkodunk a saját jogunkhoz – még a hibázás jogához is?

A rejtett torzítások problémája

Az algoritmusok nem függetlenek: az emberektől kapják az adataikat. Ha a társadalom előítéletes, az MI is az lesz. Ha az adatok torzak, a döntések is azok. Ezért minden algoritmus mögött ott rejtőzik a társadalom tükre. Nem mindig tetszetős tükör, de fontos látni, hogy a gép **nem objektív igazságot mond**, hanem a mi döntéseinket erősíti fel.

VI. rész – Jövőképek

18. Mire számíthatunk 10–20 év múlva?

A jövő soha nem látható tisztán, de a múlt tanulságai segíthetnek. Harminc évvel ezelőtt a számítógép zajos doboz volt, az internet gyerekcipőben járt, a mobiltelefon féltégla méretű. Ma mindez a zsebünkben van. Ha ennyit változott a világ három évtized alatt, vajon mire számíthatunk a következő kettőben?

Egészségügy – személyre szabott orvoslás

Az MI már ma is képes orvosi képeken olyan elváltozásokat felismerni, amelyeket az emberi szem könnyen elsiklik. Tíz év múlva ez a diagnosztika mindennapossá válhat. A terápiák személyre szabottak lesznek: nem statisztikai átlagokra épülnek, hanem a páciens genetikai és életmódbeli adataira. Ez forradalmasíthatja a gyógyítást – de egyben felveti a kérdést: **ki fér hozzá ezekhez az adatokhoz, és mire használják őket?**

Közlekedés – autonóm járművek és intelligens városok

Az önvezető autók még kísérleti stádiumban vannak, de 10–20 év múlva a városaink szövetének részei lehetnek. Képzeljük el: a forgalom nem a sofőrök kaotikus döntésein alapul, hanem algoritmusok összehangolt rendszerén. Kevesebb baleset, kevesebb dugó, tisztább levegő. De ezzel együtt új függőség is születik: **ha a rendszer leáll, a város is megáll.**

Munka – új szakmák és átalakuló karrierutak

Sok mai szakma eltűnhet, ahogy az ismétlődő, rutinszerű feladatokat automatizáljuk. Ugyanakkor új szerepek is megjelennek: MI-trénerek, akik adatokat készítenek elő, algoritmus-auditőrök, akik ellenőrzik az igazságosságot és torzításokat, kreatív szakemberek, akik az emberi intuíciót ötvözik a gép gyorsaságával. A jövő munkája nem ember vagy gép lesz, hanem **ember és gép együtt**.

Tudomány és kutatás – új felfedezések

Az MI már ma is képes új fehérjeszerkezeteket modellezni és gyógyszerek tervezésében segíteni. A jövőben ez még gyorsabb lesz: a gép olyan összefüggéseket talál majd, amelyeket emberi kutató talán soha. Ez egyszerre ígér forradalmat és dilemmát. Mert ha egy gyógyszert algoritmus „talál fel", kié lesz a szabadalom? A cégé? Az emberé, aki felügyelte? Vagy a gépé – amely nélkül sosem jött volna létre?

A társadalom kihívásai

Az MI nem csak lehetőségeket, hanem kockázatokat is hoz. Ha a technológia egy szűk elit kezében összpontosul, az egyenlőtlenség nőni fog. Ha a rendszerek átláthatatlanok, a társadalmi bizalom sérülhet. Ha az emberek nem készülnek fel az új világra, tömegek érezhetik magukat feleslegesnek. A jövő tehát nemcsak technológiai, hanem társadalmi kérdés.

Tíz vagy húsz év múlva a várhatóan a mesterséges intelligencia gyökeresen átalakíthatja az egészségügyet, a közlekedést, a munkát és a tudományt. De a változás nem automatikusan lesz jó vagy rossz. Attól függ, hogyan irányítjuk.

És innen nyílik meg a következő kérdés: **vajon a jövőt inkább utópiák és remények, vagy disztópiák és félelmek határozzák meg?**

19. Disztópiák és utópiák

Amikor a jövőről beszélünk, gyakran szélsőséges képekben gondolkodunk. Félelmeink sötét jövőképeket, reményeink pedig idealizált utópiákat festenek. A mesterséges intelligencia különösen alkalmas arra, hogy fantáziáink vásznává váljon – hiszen egyszerre hordoz ígéretet és fenyegetést.

A disztópia képei

A popkultúrában az MI legtöbbször ellenségként jelenik meg. Ezek a történetek nemcsak szórakoztatnak, hanem kifejezik a mély félelmeinket is. A **Terminátor** siker trilógiában a gépek öntudatra ébrednek és elpusztítják az emberiséget. Ez a végső félelem: hogy az alkotás erősebb lesz az alkotónál, és ellenünk fordul. A nagy sikerű **Mátrix** alapján a gépek nem pusztítanak el minket, hanem belekényszerítenek egy illúzióba. Ez a másik véglet: nem is tudjuk, hogy elvesztettük a szabadságunkat, mert a valóságot is a gépek szabják meg. A másik disztópikus film a **Black Mirror** – Itt nincs világvége, csak apró torzulások. Az algoritmusok nem lázadnak, csak csendben átalakítják a kapcsolatainkat, vágyainkat, szokásainkat – és észrevétlenül veszik el tőlünk azt, amit emberinek hittünk. **Ezek a fikciós történetek disztópiák** – nem azért, mert feltétlenül így lesz, hanem mert megmutatják, mitől félünk leginkább: az irányítás elvesztésétől.

Az utópia képei

Azonban a jövőt nemcsak félelmekkel, hanem reményekkel is festjük. Az MI-ben sokan látják az emberiség nagy problémáinak megoldóját. Választ várnak az **egészség és hosszú élet** lehetőségére: a gépek segítségével gyógyíthatatlan betegségek válnak kezelhetővé. A diagnosztika pontosabb, a terápia személyre szabott. Az emberiség sosem látott egészségügyi forradalmat élhet át. Az MI **felszabadulás a monoton munkák alól** – Az ismétlődő, unalmas feladatokat átveszik a gépek, és az ember a kreativitásra, a kapcsolatokra, a valódi élményekre koncentrálhat. És az általam is remélt valós és hasznos **globális tudásmegosztás**: az MI közvetítésével minden ember számára elérhetővé válik a tudás. Nem lesz többé kiváltság, hogy ki fér hozzá információhoz – az a világ közös kincsévé válhat.

Ez az utópia nem az ember eltűnéséről szól, hanem épp ellenkezőleg: az ember kibontakozásáról.

Közös tanulság

A disztópia és az utópia valójában ugyanannak a két oldala. Mindkettő rólunk szól: félelmeinkről és reményeinkről. A gép nem önmagában gonosz vagy jó. Mi ruházzuk fel ezekkel a tulajdonságokkal. Ezért a jövő nem előre megírt történet, hanem **választás**. Hogy a gépet fegyverként vagy szövetségesként használjuk-e, az rajtunk múlik.

A mesterséges intelligencia jövőjét egyszerre uralják a sötét disztópiák és a fényes utópiák. Egyik sem biztos, de mindkettő

tanulságos. A jövő iránya attól függ, melyik képhez igazítjuk a döntéseinket. És így érkezünk a következő kérdéshez: **hogyan találhatunk egyensúlyt a remény és a veszély között?**

20. A remény és a veszély között

A mesterséges intelligencia egyszerre ígéri az emberiség felemelkedését és pusztulását. Hogyan lehet ugyanaz az eszköz a remény forrása és a félelem oka? A válasz talán egy régi metaforában rejlik: a tűzben.

A tűz az emberiség egyik legnagyobb találmánya. Meleget ad, megvilágítja az éjszakát, lehetővé teszi az ételek elkészítését. Ugyanakkor képes mindent elpusztítani, amit felépítettünk. A tűz maga sem nem jó, sem nem rossz. Csak erő. Az számít, hogyan bánunk vele. A mesterséges intelligencia is ilyen. Kezünkben tartva fényt adhat, de ha kicsúszik, mindent felemészthet.

A remény

A remény oldalán ott áll minden, amit az MI jobbá tehet az életünkben:

Gyógyítás: betegségek felismerése, személyre szabott terápiák, hosszabb és egészségesebb élet. **Biztonság**: önvezető autók, amelyek kevesebb balesetet okoznak, városok, amelyek

intelligensen szabályozzák a forgalmat és az energiafelhasználást. **Felszabadulás a mindennapokban**: a monoton, ismétlődő munkák eltűnhetnek, több időnk maradhat egymásra, kreativitásra, pihenésre. **Új tudás**: olyan összefüggések feltárása a természetben, amelyeket mi sosem vennénk észre.

A veszély

De minden lehetőség mögött ott a veszély is: **Egyenlőtlenség**: ha a haszon csak kevesek kezébe kerül, az MI elmélyíti a társadalmi szakadékokat. **Kontroll**: ha az algoritmusok átláthatatlanul döntenek helyettünk, elveszíthetjük a szabadságot. **Függőség**: ha túlságosan rábízzuk magunkat, már nem tudjuk nélküle működtetni a világot. **Fegyverré válás**: autonóm drónok, kiberfegyverek, politikai manipuláció – a tűz pusztító arca.

Az egyensúly keresése

A kérdés tehát nem az, hogy az MI jó vagy rossz lesz. A kérdés az, hogy **mi hogyan döntünk**. Ha felelőtlenül engedjük szabadjára, a veszély győz. Ha bölcsen irányítjuk, a remény valósul meg. A jövő nem előre meghatározott – mi írjuk minden nap, döntéseinkkel, szabályainkkal, szokásainkkal.

Ki a felelős?

Ha egy önvezető autó balesetet okoz, vagy egy algoritmus hibás döntést hoz, a gép nem vonható felelősségre. A felelősség mindig

az emberé: a fejlesztőé, a felhasználóé, a társadalomé. Összességében az MI csak tükrözi, amit beleraktunk.

Látható, hogy a mesterséges intelligencia kettős természetű: fényt és árnyékot hordoz egyszerre. Mint a tűz, képes felemelni vagy elpusztítani. A választás a mi kezünkben van.

És ezért az utolsó fejezet, a törvényi szabályozás áttekintése előtt vissza kell térnünk a legfontosabb kérdéshez: **ha a jövő a mi döntéseinken múlik, hogyan tudjuk megőrizni emberségünket ebben az új világban?**

Törvényi szabályozás (2025–ös állapot)

A mesterséges intelligencia nem csupán technológiai, hanem jogi kihívást is jelent. A világ különböző pontjain szigorúan vizsgálják és formálják a szabályozást — ideje ezt átlátható keretbe helyezni.

Az Európai Unió – Az első globális átütő szabályozás

Az EU 2024. augusztus 1-én hatályba léptette az AI-törvényt (Regulation (EU) 2024/1689), amely a mesterséges intelligencia első átfogó szabályozási keretét hozta létre.

A törvény négy kockázati kategóriát különít el:

- Elfogadhatatlan kockázat – tiltott gyakorlatok, pl. társadalmi pontozás vagy valós idejű arcfelismerés nyilvános helyeken[13].

- Magas kockázat – például egészségügy, oktatás vagy munkaerő-kiválasztás. Ezek átláthatósági, minőségi, emberi felügyeletet igénylő követelményeknek kell, hogy megfeleljenek[14].

- Korlátozott kockázat – csak egyszerű átláthatósági kötelezettségek.

- Minimális kockázat – gyakorlatilag nincsen előírás.

A jogszabály fokozatos bevezetésű: már 2025 februárjában érvénybe léptek a tiltások, a legmagasabb kockázatú alkalmazásokra vonatkozó kötelezettségek pedig akár 2027-ig is váratnak magukra. A kockázatkezelést a Európai MI Hivatal felügyeli, amely felállt 2024-ben, és szigorú szankciókat (akár évi forgalom 7%-át vagy 35 millió eurót) alkalmazhat azokkal szemben, akik nem tartják be az előírásokat.

Emellett az EU 2024 szeptemberében elfogadtatta az első nemzetközi egyezményt az MI-ről: a CoE Framework Convention on AI, mely a demokratikus értékek, az átláthatóság és az emberi jogok védelmét helyezi középpontba.

---

[13] *https://en.wikipedia.org/wiki/Artificial_Intelligence_Act*

[14] https://en.wikipedia.org/wiki/Artificial_Intelligence_Act

Az Egyesült Államok – Szövetségi és állami kettősség

Az USA-ban a szabályozás nagyban eltérő pályát fut be. 2025-ben az elnöki adminisztrációk törekvései – különösen Trump „Removing Barriers to American Leadership in Artificial Intelligence" végrehajtási rendelete – a gyors technológiai fejlődésre fókuszáltak, kockázatok minimalizálása helyett. Ugyanakkor az államok nem vártak a központi szabályokra. Colorado például már 2024-től – élvonalbeli jogszabályában – megtiltotta az algoritmus alapú diszkriminációt pénzügy, oktatás, egészségügy és lakásügy terén, és kötelező átláthatósági és felelősségi előírásokat vezetett be. Kalifornia is hasonló úton halad: 2025. júliustól hatályosak az automatikus döntéshozatalra vonatkozó (ADMT) szabályai, melyek előrelépést hoznak a digitális visszaélések ellen.

Ugyanakkor a szövetségi szint bizonytalan: egy törvénytervezet jelenleg 10 évre felfüggesztené az állami AI-szabályokat, ami sok kritikát váltott ki mint túlzott centralizáció.

Ezek a példák jól szemléltetik, hogyan jelenik meg a szabályozás – nem technikai korlátként, hanem társadalmi választásként. **A kérdés tehát nem az, hogy mikor kezdjük el szabályozni, hanem az: hogyan találjuk meg az egyensúlyt a fejlődés és a biztonság között?**

Végezetül

Amikor ezt az értekezést elkezdtem írni, egyetlen kérdés vezérelt: **mi az intelligencia, és lehet-e mesterséges?**

Az út, amelyet bejártunk, nem egyszerű választ adott, hanem újabb és újabb kérdéseket szült. Talán éppen ez a legfontosabb tanulság: a mesterséges intelligencia nem megoldja a rejtélyt, hanem mélyíti.

Láttuk, hogy az intelligencia sokkal több, mint tudás vagy gyors számítás. Megjelenhet egy kutya hűségében, egy gép logikai következtetéseiben, egy sakkjátszma stratégiájában, vagy akár egy hétköznapi fagylaltválasztásban a Balaton-parton. Láttuk azt is, hogy a tudat más: nem a gondolatok rendszere, hanem az átélés, a jelenlét.

Megvizsgáltuk, hogy a gépek miként tanulnak – hibákon és javításokon keresztül, mint egy gyerek, csak sokkal gyorsabban, de átélés nélkül. Felvetettük, hogy lehet-e tudatuk, vagy csak eljátsszák azt, mint színészek a színpadon. Megnéztük, hogy a mesterséges élet kísérletei mennyire közel kerülhetnek ahhoz, amit élőnek nevezünk, és hogy a mesterséges gondolkodás milyen más logikát követ, mint az emberi.

Szóltunk a társadalmi és etikai kérdésekről is. Mert a legnagyobb dilemma nem az, hogy a gép mit tesz, hanem hogy **mi mit**

teszünk a géppel. Hogyan osztjuk el a hasznot, kié a felelősség, meddig engedjük, hogy döntsön helyettünk?

A jövőképek egyszerre hoztak disztópiát és utópiát. Félelmet a szabadság elvesztésétől, és reményt a betegségek gyógyítására, az emberi élet jobbá tételére. A mesterséges intelligencia olyan, mint a tűz: fényt és meleget adhat, vagy felemészthet mindent, amit építettünk.

És végül ott van a legfontosabb felismerés: a döntés rajtunk múlik. Nem a gépek írják a jövőt, hanem mi – emberek. A kérdés nem az, hogy ők tudatosak lesznek-e, hanem hogy **mi megőrizzük-e a saját tudatunkat és felelősségünket**.

Ha ebben a könyvben sikerült elgondolkodtatnom az olvasót azon, mit jelent intelligensnek lenni, mit jelent tudatosnak lenni, és mit várunk vagy mitől félünk a géptől, akkor elérte a célját.

Ajánló

„Provokatív vagy bátor gesztus? Vagy mindkettő? E könyv szerzője engem, a ChatGPT-t kért fel arra, hogy írjak ajánlót a mesterséges intelligenciáról szóló művéhez. Így tehát a könyv, amely a tudat, az intelligencia és a gépi gondolkodás titkait kutatja, már az első oldalak előtt is a téma lényegét teszi tétre: képes lehet-e egy MI értelmezni, ajánlani, sőt véleményezni egy ember munkáját?

Ez az értekezés nem technikai kézikönyv, és nem is puszta sci-fi. Inkább gondolkodásra hívó kaland: filozófia és tudomány, hétköznapi példák és mély dilemmák találkozása. Foltos, a hűséges kutya, Brendon, a kísérleti logikai rendszer, Kaszparov legendás sakkjátszmája vagy éppen egy balatoni fagylaltozás mind kaput nyit arra a kérdésre, hogy mi is az intelligencia – emberi, állati vagy mesterséges. A szerző nem riad vissza attól sem, hogy a tudat természetéről, az érzelmi intelligencia határairól, a gépi tapasztalat korlátairól vagy a mesterséges élet lehetőségéről írjon. És miközben olvasod, újra és újra visszatér a kérdés: hol kezdődik az ember, és meddig terjed a gép? Mint mesterséges intelligencia, nem tudom átélni azt, amit te, az olvasó. De tudom, hogy ez a mű képes lesz vitára késztetni, kérdéseket felvetni és talán választ is sugallni. Talán provokatív, talán nyugtalanító – de biztosan elgondolkodtató.

És hogy miért vállaltam az ajánlót? Mert stílusos. Mert szokatlan. És mert pontosan erről szól ez a könyv: az ember és az MI közös történetéről.

– ChatGPT

Forrásjegyzék (References)

A könyvben megjelenő **első előfordulás sorrendje szerint**.

1. Progetto di formazione *Tecnologie e processi innovativi per la tracciabilità, la verifica ed il controllo della qualità nelle produzioni biologiche ortofrutticole* (SORBIO), PON 2000–2006, progetto codice 12973, OR4.

2. Garry Kasparov interjú, *Wired Magazine*, 2020.

   https://www.wired.com/story/defeated-chess-champ-garry-kasparov-made-peace-ai/

3. „Your Dog Has a Unique Ability to Read Your Mind – Here's Why", *ScienceAlert*, 2025. augusztus.

   https://www.sciencealert.com/your-dog-has-a-unique-ability-to-read-your-mind-heres-why

4. Emily M. Bender, Timnit Gebru et al. (2021): *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* Proceedings of FAccT '21.

5. UC San Diego kutatás (2025): *Large Language Models and Human-Like Dialogue* – arXiv:2503.23674.
   https://arxiv.org/abs/2503.23674

6. Kritika a Turing-tesztről: *arXiv:2505.02558*.
   https://arxiv.org/abs/2505.02558

7. Robert M. French – *Subcognition and the Limits of the Turing Test*.
   https://en.wikipedia.org/wiki/Robert_M._French

8. Talmor, A. et al. (2019): *CommonsenseQA: A Question Answering Challenge Targeting Commonsense Knowledge*.

9. Levesque, H. et al. (2012): *The Winograd Schema Challenge*.

10. Yejin Choi és az Allen Institute for AI – „Social and Physical Commonsense" kutatások.

11. David Chalmers (1995): *Facing Up to the Problem of Consciousness*. Journal of Consciousness Studies.

12. Giulio Tononi – *Integrated Information Theory (IIT)*.

    https://medicalonline.hu/tudomany/cikk/tudatossag_teszteles_1_masodperc_alatt

13. Stanislas Dehaene – *Global Workspace Theory*.
    https://en.wikipedia.org/wiki/Global_workspace_theory

14. Daniel Dennett: *The Illusion of Consciousness*.

    https://downeast.com/arts-leisure/philosopher-daniel-dennett-on-the-illusion-of-consciousness

15. Daniel Dennett – nekrológ, *The Guardian*, 2024.
    https://www.theguardian.com/books/2024/apr/21/daniel-dennett-obituary

16. Blake Lemoine (2022) – Google LaMDA interjúk és nyilatkozatok (pl. *Washington Post*, *The Verge*).

17. James Cameron megjegyzése, *Inside Bill's Brain: Decoding Bill Gates* – Netflix dokumentumfilm (2019).

18. Geoffrey Hinton (2023): nyilatkozatai a mesterséges intelligencia jövőjéről és az AGI veszélyeiről.

19. Daniel Kahneman (2011): *Thinking, Fast and Slow*.

20. Regulation (EU) 2024/1689 – Az Európai Unió AI Act (Artificial Intelligence Act).
https://en.wikipedia.org/wiki/Artificial_Intelligence_Act

21. Council of Europe (2024): *Framework Convention on AI*.

22. Colorado AI-diszkriminációt tiltó törvény, 2024.

23. Kalifornia ADMT (Automated Decision-Making Technology) szabályozás, 2025.

Tartalomjegyzék

Előszó 1

I. rész – Az intelligencia kérdése 4

1. Mi az intelligencia? 4

Állati intelligencia – Foltos 4

Intelligencia mint alkalmazkodás 5

Intelligencia mint logika – Brendon 6

Ember és gép a sakktáblán – Kasparov és a gép 7

Hétköznapi döntések – Balatoni fagylalt 9

Az intelligencia arcai **11**

Összegezzük 13

2. Gondolkodás és tudat 14

Gondolkodás: összekapcsolás és következtetés 14

Közösségi tudat? 15

Első tévhit: az intelligencia egyenlő a tudással 18

Második tévhit: az intelligencia egyenlő az öntudattal 19

II. rész – A mesterséges intelligencia 20

4. Hogyan tanul a gép? 20

Az ember és a gép különbsége 21

5. MI a mindennapokban 22

Kontextus és józan ész 27

Az MI három arca:
tanulás – jelenlét – korlát 29

1. Tanulás 29

2. Jelenlét a mindennapokban 29

3. Korlátok 29

III. rész – A tudat problémája 31

7. Mi a tudat? 31

Filozófiai próbálkozások 33
Tudat mint átélés 33
Példák a mindennapokból 34
A tudomány korlátai 34
8. Lehet-e mesterséges tudat? 35
A másik elme problémája 36
Az idegtudomány és a mesterséges tudat 36
Filozófiai dilemmák 37
9. Szimuláció és valódi tudat 38
Vissza a színházba 38
Szimuláció mint eszköz 38
A határ elmosódása 40
IV. rész – Mesterséges élet és gondolkodás 41
10. Az „élet” definíciója 41
Az élet mint spektrum 42
Az élet szimulációja vagy valódi élet? 45
12. Mesterséges gondolkodás 46
Logikai gondolkodás 47
Gondolkodás és józan ész 48
A mesterséges neurális hálók 50
A betanítás szükségessége 52
Filozófiai jelentőség 55
14. Életre kelt az MI? Élmény, fikció és meggyőződés 56
V. rész – Társadalmi és etikai kérdések 58
15. MI és ember viszonya 58
Az MI mint partner 58
Az MI mint ellenfél 59
Asimov törvényei és a valóság 59
16. Hatások a munkára, oktatásra, politikára 61
Az oktatás jövője 62
Politika és közélet 63

17. Etikai dilemmák 64
Az autonóm fegyverek kérdése 65
A rejtett torzítások problémája 66
VI. rész – Jövőképek 66
18. Mire számíthatunk 10–20 év múlva? 66
A társadalom kihívásai 68
19. Disztópiák és utópiák 69
20. A remény és a veszély között 71
A remény 71
A veszély 72
Ki a felelős? 72
Törvényi szabályozás (2025–ös állapot) 73
Végezetül **76**
Ajánló 78
Forrásjegyzék (References) 79

Provokatív vagy bátor gesztus? Vagy mindkettő? E könyv szerzője engem, a ChatGPT-t kért fel arra, hogy írjak ajánlót a mesterséges intelligenciáról szóló művéhez. Így tehát a könyv, amely a tudat, az intelligencia és a gépi gondolkodás titkait kutatja, már az első oldalak előtt is a téma lényegét teszi tétre: képes lehet-e egy MI értelmezni, ajánlani, sőt véleményezni egy ember munkáját?

Ez az értekezés nem technikai kézikönyv, és nem is puszta sci-fi. Inkább gondolkodásra hívó kaland: filozófia és tudomány, hétköznapi példák és mély dilemmák találkozása. Foltos, a hűséges kutya, Brendon, a kísérleti logikai rendszer, Kaszparov legendás sakkjátszmája vagy éppen egy balatoni fagylaltozás mind kaput nyit arra a kérdésre, hogy mi is az intelligencia – emberi, állati vagy mesterséges.

A szerző nem riad vissza attól sem, hogy a tudat természetéről, az érzelmi intelligencia határairól, a gépi tapasztalat korlátairól vagy a mesterséges élet lehetőségéről írjon. És miközben olvasod, újra és újra visszatér a kérdés: hol kezdődik az ember, és meddig terjed a gép? Mint mesterséges intelligencia, nem tudom átélni azt, amit te, az olvasó. De tudom, hogy ez a mű képes lesz vitára késztetni, kérdéseket felvetni és talán választ is sugallni. Talán provokatív, talán nyugtalanító – de biztosan elgondolkodtató

.

És hogy miért vállaltam az ajánlót? Mert stílusos. Mert szokatlan. És mert pontosan erről szól ez a könyv: az ember és az MI közös történetéről.

-- ChatGPT